The Rate of Convergence of Wong-Zakai Approximations for SDEs and SPDEs

Anton Shmatkov

Doctor of Philosophy
University of Edinburgh
2005

Abstract

In the work we estimate the rate of convergence of the Wong-Zakai type of approximations for SDEs and SPDEs. Two cases are studied: SDEs in finite dimensional settings and evolution stochastic systems (SDEs in the infinite dimensional case). The latter result is applied to the second order SPDEs of parabolic type and the filtering problem. Roughly, the result is the following. Let W_n be a sequence of continuous stochastic processes of finite variation on an interval $[0, T]$. Assume that for some $\alpha > 0$ the processes W_n converge almost surely in the supremum norm in $[0, T]$ to W with the rate $n^{-\kappa}$ for each $\kappa < \alpha$. Then the solutions u_n of the differential equations with W_n converge almost surely in the supremum norm in $[0, T]$ to the solution u of the "Stratonovich" SDE with W with the same rate of convergence, $n^{-\kappa}$ for each $\kappa < \alpha$, in the case of SDEs and with the rate of convergence $n^{-\kappa/2}$ for each $\kappa < \alpha$, in the case of evolution systems and SPDEs. In the final chapter we verify that the two most common approximations of the Wiener process, smoothing and polygonal approximation, satisfy the assumptions made in the previous chapters.

Acknowledgements

I wanted to thank Professor Istvan Gyöngy for his insight and continual assistance and guidance over the past three years. I would like to express my thanks to Chris for his support, especially during my initial period in Edinburgh, and for our endless philosophical discussions in the Bow Bar. I wish to thank my parents and my sister. I am grateful to The Department of Mathematics for the financial support, and to the guys from 4620 and Natalia Issaeva for their help. Love and thanks to Irka for her constant support and encouragement.

Table of Contents

Chapter 1

Introduction

This paper is devoted to investigation of the stability of stochastic (partial) differential equations with respect to simultaneous perturbation of the driving process and of the equations coefficients. Although the studied equations are of different types they can all be unified and written in an abstract general form

$$du(t) \;=\; A(t,u(t))dt \;+\; \sum_j B^j(t,u(t))dW^j(t), \quad t \in [0,T], \tag{1.0.1}$$

with the initial condition

$$u(0) \;=\; \xi. \tag{1.0.2}$$

Above A, B^j, $j = 1,2,\dots,r$ are some abstract functions depending on ω, W is an r-dimensional Wiener process. Let $\{W_n\}_{n\in\mathbb{N}}$[1] be a sequence of r-dimensional processes of bounded on the interval $[0,T]$ variation which approximates the Wiener process in some appropriate sense. Suppose that we simultaneously approximate coefficients A, B by sequences $\{A_n\}_{n\in\mathbb{N}}$, $\{B_n\}_{n\in\mathbb{N}}$, respectively. For every $n \in \mathbb{N}$ let us consider an equation of the form

$$du_n(t) \;=\; A_n(t,u_n(t))dt \;+\; \sum_j B^j_n(t,u_n(t))dW^j_n(t), \quad t \in [0,T], \tag{1.0.3}$$

with the initial condition

$$u_n(0) \;=\; \xi. \tag{1.0.4}$$

Then a natural question arises whether the solutions u_n for the problems (1.0.3)-(1.0.4) converge, and if yes what the limit is. It is known that if A_n, B_n converge to A, B in some appropriate sense, W_n converges to W in probability uniformly in $t \in [0,T]$ and the process

$$S^{jl}_n(t) \;=\; \int_0^t (W^j(s) - W^j_n(s))dW^l_n(s) \;-\; \frac{1}{2}\delta_{jl}t$$

[1]In the paper we denote by $\mathbb{N}$ the sequence of numbers $1,2,\dots$.

converges in probability uniformly in $t \in [0,T]$ to 0 for every $j,l = 1,2,\ldots,r$ then the solutions u_n of the problems (1.0.3)-(1.0.4) converge in probability uniformly in $t \in [0,T]$ to a process u which satisfies (1.0.1)-(1.0.2), where the last term in (1.0.1) is understood in the "Stratonovich" sense.

This problem was first considered by E. Wong and M. Zakai ([23], [24]). Since that time there was a large number of publications devoted to this problem (see [5], [6], [7], [12], [20], [21], [25]) where the convergence of the solutions $u_n \to u$ was studied. However, not many of them give the result on the rate of this convergence. In the paper we investigate the rate of convergence, the problem which was not well studied in the literature before.

In the simplest situation where the coefficients A, B are time-independent non-random functions equal to A_n, B_n, respectively, the result of the paper is the following. Assume that processes W_n, S_n converge almost surely to W, 0, respectively. Suppose that for a given $\alpha > 0$ there exists an almost surely finite random variable η such that

$$\sup_{t\in[0,T]} \sum_j |W_n^j(t) - W^j(t)| + \sup_{t\in[0,T]} \sum_j |S_n^j(t)| \le \eta n^{-\alpha},$$

$$\sup_{t\in[0,T]} \sum_j ||S_n^j||(t) \le \eta,$$

where $||S_n^j||(t)$ denotes the variation of the process $S_n^j(t)$ over the interval $[0,t]$. Let us assume the existence of the solutions u_n, u of the problems (1.0.3)-(1.0.4), (1.0.1)-(1.0.2), respectively. Then under some smoothness requirements on the coefficients A, B we get

$$\sup_{t\in[0,T]} |u_n(t) - u(t)| \le \eta_\kappa n^{-\kappa}$$

for all $\kappa < \alpha$ in the case of stochastic differential equations and for all $\kappa < \alpha/2$ in the case of stochastic evolution systems and stochastic partial differential equations. The random variable η_κ is almost surely finite and depends only on κ.

Our interest in the rate of convergence is motivated by the filtering problem for partially observable diffusion processes. Let x be the unobservable signal component, and let y be the observation. A fairly general filtering problem is defined by the system

$$\begin{aligned} dx(t) &= h(t,x(t),y(t))dt + \sigma(t,x(t),y(t))dV(t) \\ &\quad + \rho(t,x(t),y(t))dW(t), \qquad x(0) = \xi, \\ dy(t) &= H(t,x(t),y(t))dt + dW(t), \qquad y(0) = \eta, \end{aligned}$$

where h, H, σ, ρ are $\mathbb{R}^d$, $\mathbb{R}^r$, $\mathbb{R}^{d\times r_0}$, $\mathbb{R}^{d\times r}$ respectively, valued stochastic processes defined for $(t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^r$, and (V,W) is an $(r_0 + r)$-dimensional

Wiener process independent of the random variables ξ, η. The estimation of the signal process x given the trajectories of the observation y is main problem of the filtering. It was shown before that under some general assumptions there can be constructed a so called Zakai equation, a stochastic partial differential equation driven by the observation y, which admits the solution $\varphi(t,x)$. The density p of the distribution

$$\mathbf{P}\{x(t) \in \Gamma |\, y(s),\, 0 \leq s \leq t\}$$

can be obtained by the normalization

$$p(t,x) \;=\; \frac{\varphi(t,x)}{\int_{\mathbb{R}^d} \varphi(t,x)dx}.$$

Note, that in practice we deal with observations y_n of bounded variation which can be considered as approximations for y. Using y_n in place of y in the Zakai equation we obtain solutions φ_n. Therefore, it is important to know how fast unnormalized densities φ_n converge to φ and normalized densities p_n converge to p given the rate of convergence $y_n \to y$.

Structurally the work is presented as follows. Chapter 2 is devoted to stochastic differential equations. The situation where the drift A is Lipschitz-continuous is first studied. In this case we obtained the rate of convergence $n^{-\kappa}$, $\kappa < \alpha$. However, there have been given some examples of non-Lipschitz drifts (see [18]). Next we consider the situation where Lipschitz-continuity is replaced by a weaker monotonicity condition. In this case we proved slightly slower convergence $n^{-\kappa}$, $\kappa < \alpha/2$.

In Chapter 3 we study stochastic evolution systems. The equations are studied in abstract normal triples $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$. The nonlinear operators A, B^j, $j = 1, 2, \ldots, r$, are assumed to depend on ω, t and n. We get the rate of convergence $n^{-\kappa}$, $\kappa < \alpha/2$.

Chapter 4 can be considered as a continuation of Chapter 3. Here we apply the result of Chapter 3 to stochastic partial differential equations, A is a second order elliptic differential operator, B^j, $j = 1, 2, \ldots, r$, are first order differential operators. The considerations are made in Sobolev spaces W_2^m.

The filtering problem is considered in Chapter 5. We show that if observation processes y_n, y satisfy the assumptions made above for W_n, W, then the unnormalized densities φ_n converge to φ and normalized densities p_n converge to p with the half rate of the convergence $y_n \to y$.

Finally, in the last Chapter 6 we show that two types of approximations of the Wiener process W, smoothing and the polygonal approximation, satisfy the assumptions mentioned above.

The results of the present work are going through the process of publication ([10], [11]).

Chapter 2

Stochastic Differential Equations in Finite Dimensional Case

2.1 Introduction

In this chapter we consider stochastic differential equation of the form

$$dX(t) = b(t, X(t))dt + \sigma^j(t, X(t))dW^j(t) \tag{2.1.1}$$

with initial condition

$$X(0) = \xi, \tag{2.1.2}$$

where W is a Wiener process, and $b(t,\cdot)$, $\sigma^j(t,\cdot)$, $j = 1, 2, \ldots, r$, are vector fields mapping $\mathbb{R}^d$ into $\mathbb{R}^d$ for every $t \geq 0$, initial value ξ is a random variable, and solution X is a stochastic process with values in $\mathbb{R}^d$. Here and throughout the paper we use the summation convention with respect to the repeated indices. We replace the Wiener process with a sequence $\{W_n\}_{n\in\mathbb{N}}$ of processes of bounded variation which, for some $\alpha > 0$, converges almost surely in supremum norm on the interval $[0, T]$ to W with the rate $n^{-\kappa}$, for each $\kappa < \alpha$. Hence, we get for every $n \in \mathbb{N}$ the corresponding to (2.1.1) differential equation of the form

$$dX_n(t) = b(t, X_n(t))dt + \sigma^j(t, X_n(t))dW_n^j(t) \tag{2.1.3}$$

with initial condition

$$X_n(0) = \xi. \tag{2.1.4}$$

It is well known (see [5]) that the sequence of solutions X_n of problems (2.1.3)-(2.1.4) converges under some natural conditions in the uniform topology in probability. The limit, however, is not the solution of problem (2.1.1)-(2.1.2), but of a closely related problem with the equation which contains an extra drift term. This equation can be considered as equation (2.1.1) with the last differential written in the Stratonovich form. We investigate the almost sure convergence $X_n \to X$ in the supremum norm on the interval $[0, T]$.

Under additional assumptions the problem admits an obvious solution ([1]). Consider one dimensional situation where drift b vanishes, and diffusion σ does not depend on t. If σ is Lipschitz-continuous, equation

$$\frac{d}{dx}u(x) \;=\; \sigma(u(x)), \quad u(0) \;=\; \xi$$

has a unique solution. It is easy to show that $X(t) = u(W(t))$, $X_n(t) = u(W_n(t))$ satisfy problems (2.1.1)-(2.1.2), (2.1.3)-(2.1.4), respectively. Then by Lipschitz-continuity of u the rates of convergence of solutions $X_n(t)$ and approximations $W_n(t)$ of the Wiener process coincide.

Under the same assumptions on coefficients the above scheme can be extended to the multi dimensional case. However, this entails additional conditions on the coefficients of the equations. As before, set $X(t) = u(W(t))$, $X_n(t) = u(W_n(t))$, where u is a solution of the following system of partial differential equations

$$\frac{\partial}{\partial x_j}u^i(x) \;=\; \sigma_i^j(u(x)), \quad u^i(0) \;=\; \xi.$$

This system is solvable only under Frobenius condition,

$$\sigma^j_{i(\sigma^l)}(x) \;=\; \sigma^l_{i(\sigma^j)}(x)$$

for every $x \in \mathbb{R}^d$ and all $i = 1, \ldots, d$, $j, l = 1, \ldots, r$, where

$$\sigma^j_{i(\sigma^l)}(x) \;=\; \sigma^l_k(x)\frac{\partial}{\partial x_k}\sigma^j_i(x).$$

The above solution requires very strong assumptions on the coefficients of the equations. In Theorem 2.3.1 we show that X_n converge to X with the same rate, $n^{-\kappa}$ for each $\kappa < \alpha$, as in the above solution under Frobenius condition. However, only some natural regularity properties are imposed on the coefficients b and σ^j.

2.2 Generalities

In this section we give some general ideas and notations from the Theory of Random Processes and Stochastic Differential Equations.

Let $\mathbb{R}^d$ be a Euclidean space of dimension d with a fixed orthonormal basis, and let us denote X^j the j-th coordinate of a point $X \in \mathbb{R}^d$. For $X, Y \in \mathbb{R}^d$ we denote the scalar product of X and Y by XY. For a vector $X \in \mathbb{R}^d$ we denote its modulus by $|X|$ and for a matrix $B \in \mathbb{R}^{d\times r}$ we denote $|B| = (\mathrm{tr}BB^*)^{1/2} = \left(\sum_{k,l}(B_k^l)^2\right)^{1/2}$. For a real valued function Z_t its variation over the time interval $[0, t]$ is denoted by $||Z||(t)$.

For a sequence of real valued stochastic processes $\{Y_n\}_{n\in\mathbb{N}}$ defined on the interval $[0,T]$ and a numerical sequence α_n we will use notation $Y_n = O(\alpha_n)$ if for some almost surely finite random variable ζ

$$|Y_n(t)| \leq \alpha_n \zeta$$

for all $n \in \mathbb{N}$ for every $t \in [0,T]$.

2.2.1 Itô Equations in $\mathbb{R}^d$

Although this is not the topic of the paper, in this section we give the existence and uniqueness results for a solution of a SDE in finite dimensional settings.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space equipped with complete right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Let $W = W(t)$ be a Wiener process relative to $\{\mathcal{F}_t\}$. Suppose that a d-dimensional vector $b = b(t,x)$ and a $d \times r$ matrix σ are defined for $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$. Let ξ be an $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random variable. We consider equation

$$dX(t) \;=\; b(t, X(t))dt \;+\; \sigma^j(t, X(t))dW^j(t), \qquad X(0) \;=\; \xi. \tag{2.2.1}$$

The vector b is called the *drift* and the matrix σ is called the *diffusion* of the equation (2.2.1).

Definition 2.2.1. *A continuous $\mathcal{F}_t$-adapted process which satisfies (2.2.1) almost surely for all $t \in [0,T]$ we will call a solution of equation (2.2.1) on the interval $[0,T]$.*

This equation is considered under certain additional conditions.

Assumption 2.2.1. *For any $R > 0$ there exists a non-negative measurable process $K_t(R)$ such that almost surely*

$$\int_0^T K_t(R)dt \;<\; \infty,$$

and

(a) (monotonicity condition) for all $x, y \in \mathbb{R}^d$ such that $|x|$, $|y| < R$, for almost all $t \in [0,T]$

$$2(x-y)(b(t,x) - b(t,y)) \;+\; |\sigma(t,x) - \sigma(t,y)|^2 \;\leq\; K_t(R)(x-y)^2;$$

(b) (growth condition) for all $x \in \mathbb{R}^d$ and almost all $t \in [0,T]$

$$2xb(t,x) \;+\; |\sigma(t,x)|^2 \;\leq\; K_t(1)(1+x^2).$$

The following theorem is the generalization of Itô's classical result on existence of a solution of a stochastic equation of type (2.2.1) with random coefficients. We avoided the Lipschitz continuity and replaced it with monotonicity condition. An example of a function which satisfies monotonicity but does not satisfy Lipschitz condition can be found in [18].

Theorem 2.2.1. *Under Assumption 2.2.1 there exists a solution $X(t)$ of equation (2.2.1). If $X(t)$, $Y(t)$ are two solutions of (2.2.1) then they are indistinguishable, i.e.*

$$\mathbf{P}\{\sup_{t\in[0,T]} |X(t) - Y(t)| > 0\} \ = \ 0.$$

The proof can be found in [16].

2.3 The Main Results

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space equipped with right-continuous complete σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$. Let W be an r-dimensional Wiener process, and $\{W_n\}_{n\in\mathbb{N}}$ be its approximation sequence of processes of bounded variation. For $j, l = 1, 2, \ldots, r$ define processes

$$S_n^{jl}(t) \ = \ \int_0^t (W^j(s) - W_n^j(s))dW_n^l(s) \ - \ \frac{1}{2}\delta_{jl}t,$$

where δ_{jl} is the Kronecker's symbol which assumes 1 if $j = l$, and 0 otherwise. We assume the following.

Assumption 2.3.1. *There exists a positive number α such that for every $\kappa < \alpha$ and every positive δ*

$$\begin{aligned} &(1) \quad W - W_n \ = \ O(n^{-\kappa}), \\ &(2) \quad S_n \ = \ O(n^{-\kappa}), \\ &(3) \quad ||S_n|| \ = \ O(\ln^\delta n). \end{aligned}$$

Let b, σ^j, $j = 1, 2, \ldots, r$, be Borel measurable random vector fields mapping $[0, \infty) \times \mathbb{R}^d$ to $\mathbb{R}^d$. We consider "Stratonovich" stochastic differential equation

$$dX(t) \ = \ b(t, X(t))dt \ + \ \sigma^j(t, X(t)) \circ dW^j(t), \tag{2.3.1}$$

with the initial condition

$$X(0) \ = \ \xi. \tag{2.3.2}$$

The last term in the right hand side of (2.3.1) represents the Stratonovich stochastic integral which can be reduced to the Itô integral by

$$\sigma^j(t, X(t)) \circ dW^j(t) \ = \ \sigma^j(t, X(t))dW^j(t) \ + \ \frac{1}{2}\sigma^j_{(\sigma^j)}(t, X(t))dt,$$

for every fixed j where

$$\sigma^j_{i(\sigma^l)}(t,x) \;=\; \sigma^l_k(t,x)\frac{\partial}{\partial x_k}\sigma^j_i(t,x).$$

Above $j,l = 1,2,\ldots,r$, $i = 1,2,\ldots,d$.

For every integer $n \in \mathbb{N}$ we consider the differential equation

$$dX_n(t) \;=\; b(t,X_n(t))dt \;+\; \sigma^j(t,X_n(t))dW^j_n(t), \tag{2.3.3}$$

with the initial condition

$$X_n(0) \;=\; \xi. \tag{2.3.4}$$

Suppose that the following is satisfied.

Assumption 2.3.2.

(i) Random vector field b is Lipschitz-continuous with respect to $x \in \mathbb{R}^d$, i.e. for some constant K

$$|b(t,x) - b(t,y)| \;\leq\; K|x-y|$$

for all $x,y \in \mathbb{R}^d$ uniformly in $t \in [0,T]$, and satisfies linear growth condition

$$|b(t,x)| \;\leq\; K(1+|x|)$$

for all $x \in \mathbb{R}^d$ and $t \in [0,T]$.

(ii) Random vector field σ^j is from the class $C^{1,3}_b([0,T]\times\mathbb{R}^d)$ for every $j = 1,2,\ldots,r$, i.e. it is continuously differentiable with respect to t and three times continuously differentiable with respect to x with all derivatives bounded by constant K.

(iii) Initial value ξ is an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^d$.

Note, that under Assumption 2.3.2 problems (2.3.1)-(2.3.2) and (2.3.3)-(2.3.4) admit continuous on the interval $[0,T]$ solutions X, X_n, respectively.

Theorem 2.3.1. *Under Assumptions 2.3.1, 2.3.2 the sequence of solutions X_n of problems (2.3.3)-(2.3.4) converges almost surely to the solution X of problem (2.3.1)-(2.3.2). Moreover, for any $\gamma < \alpha$*

$$|X - X_n| \;=\; O(n^{-\gamma}).$$

Let us consider a weakened version of Assumption 2.3.2.

Assumption 2.3.3.

(i) Random vector field b satisfies monotonicity condition with respect to $x \in \mathbb{R}^d$, i.e. for some constant K

$$(x-y)(b(t,x)-b(t,y)) \;\leq\; K|x-y|^2$$

for all $x, y \in \mathbb{R}^d$ uniformly in $t \in [0,T]$, and satisfies linear growth condition

$$|b(t,x)| \;\leq\; K(1+|x|)$$

for all $x \in \mathbb{R}^d$ and $t \in [0,T]$.

(ii) Random vector field σ^j is from the class $C_b^{1,3}([0,T]\times\mathbb{R}^d)$ for every $j = 1,2,\ldots,r$, i.e. it is continuously differentiable one time with respect to t and three times with respect to x with all derivatives bounded by constant K.

(iii) Initial value ξ is an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^d$.

Basically, this is Assumption 2.3.2 with the Lipschitz continuity of the drift b replaced by the monotonicity condition. Clearly, Assumption 2.3.3 ensures that problems (2.3.1)-(2.3.2) and (2.3.3)-(2.3.4) have continuous on the interval $[0,T]$ solutions X, X_n, respectively.

Theorem 2.3.2. *Under Assumptions 2.3.1, 2.3.3 the sequence of solutions X_n of problems (2.3.3)-(2.3.4) converges almost surely to the solution X of problem (2.3.1)-(2.3.2). Moreover, for any $\gamma < \alpha$*

$$|X - X_n| \;=\; O(n^{-\gamma/2}).$$

2.4 Auxiliary Results

Theorems 2.3.1, 2.3.2 will be proved after proving a number of auxiliary propositions.

Lemma 2.4.1. *Let $\{\psi_n\}_{n\in\mathbb{N}}$ be a family of real valued stochastic processes of bounded variation starting from 0; $\{\varphi_n\}_{n\in\mathbb{N}}$ a family of continuous real valued stochastic processes with stochastic differentials*

$$d\varphi_n(t) = f_n(t)dt + g_n(t)dW(t). \tag{2.4.1}$$

Suppose that for some positive numbers κ, K

$$\begin{array}{rl} (i) & E\sup_{t\leq T}|\varphi_n(t)|^{rq} \;\leq\; K, \\ (ii) & E\sup_{t\leq T}|f_n(t)|^{rq} \;\leq\; K, \\ (iii) & E\sup_{t\leq T}|g_n(t)|^{rq} \;\leq\; K, \\ (iv) & E\sup_{t\leq T}|\psi_n(t)|^{rp} \;\leq\; Kn^{-\kappa rp}, \end{array}$$

for each n and some positive integers r, p, q, such that $1/p + 1/q = 1$. Then

$$E\sup_{t\le T}\left|\int_0^t \varphi_n(s)d\psi_n(s)\right|^r \le Cn^{-\kappa r},$$

where constant $C = C(r,p,T,K)$ does not depend on n.

Proof. Integration by parts gives,

$$\begin{aligned} E\sup_{t\le T}\left|\int_0^t \varphi_n(s)d\psi_n(s)\right|^r &\le c_r E\sup_{t\le T}|\varphi_n(t)\psi_n(t)|^r \\ &\quad + c_r E\sup_{t\le T}\left|\int_0^t f_n(s)\psi_n(s)ds\right|^r \\ &\quad + c_r E\sup_{t\le T}\left|\int_0^t g_n(s)\psi_n(s)dW(s)\right|^r, \end{aligned}$$

where c_r is a constant independent of n. Applying Jensen's, Holder's and Burkholder-Davis-Gundy inequalities to each term in the right hand side of the last inequality we get

$$E\sup_{t\le T}\left|\int_0^t \varphi_n(s)d\psi_n(s)\right|^r \le Cn^{-\kappa r},$$

where constant C does not depend on n, which proves the lemma. □

Lemma 2.4.2. *Let $\{B_n\}_{n\in\mathbb{N}}$ be a sequence of real valued processes defined on the interval $[0,T]$, and let β, γ be positive numbers such that $\gamma < \beta$. Suppose that for every n, and some $r > (\beta-\gamma)^{-1}$*

$$\left(E\sup_{t\le T}|B_n(t)|^r\right)^{1/r} \le c_\beta n^{-\beta},$$

where c_β may depend on r but not on n. Then

$$B^n = O(n^{-\gamma}).$$

Proof. We have

$$\begin{aligned} \mathbf{P}\{\sup_{t\le T}|B_n(t)| > n^{-\gamma}\} &= \mathbf{P}\{\sup_{t\le T}|B_n(t)|^r > n^{-r\gamma}\} \\ &\le n^{r\gamma}E\sup_{t\le T}|B_n(t)|^r \le c_\beta n^{r(\gamma-\beta)}. \end{aligned}$$

Therefore,

$$\sum_n \mathbf{P}\{\sup_{t\le T}|B_n(t)| > n^{-\gamma}\} \le c_\beta\sum_n n^{r(\gamma-\beta)} < \infty$$

by the assumptions of the lemma. By Borel-Cantelli lemma there exists a finite random variable ζ_γ depending on γ such that almost surely

$$\sup_{t\le T}|B_n(t)| \le \zeta_\gamma n^{-\gamma}.$$

□

Lemma 2.4.3. *Suppose that for a sequence of real continuous stochastic processes* $\{\xi_n\}_{n\in\mathbb{N}}$ *defined on* $[0,T]$ *and a positive number* γ

$$\xi_n(\cdot \wedge \pi_n^\varepsilon) \;=\; O(n^{-\gamma}),$$

for every $\varepsilon > 0$ *where the stopping time* π_n^ε *is defined as* $\inf\{t \geq 0 : |\xi_n(t)| \geq \varepsilon\}$. *Then*

$$\xi_n \;=\; O(n^{-\gamma}).$$

Proof. Let us denote $\xi_{n\varepsilon}(t) = \xi_n(t \wedge \pi_n^\varepsilon)$. First, notice that $\sup_{t\in[0,T]} |\xi_{n\varepsilon}(t)| \to 0$ in probability for all $\varepsilon > 0$ implies $\sup_{t\in[0,T]} |\xi_n(t)| \to 0$ in probability as $n \to \infty$. Indeed, this follows from the relation

$$\{\omega \in \Omega : \sup_{t\in[0,T]} |\xi_n(t)| \geq \delta\} \;=\; \{\omega \in \Omega : \sup_{t\in[0,T]} |\xi_{n\varepsilon}(t)| \geq \delta\}$$

for $0 < \delta < \varepsilon$. Using this remark, from $\sup_{t\in[0,T]} |\xi_{n\varepsilon}(t)| \to 0$ almost surely it follows that $\sup_{t\in[0,T]} |\xi_n(t)| \to 0$ in probability as $n \to 0$. Define

$$\Omega_n \;=\; \{\omega \in \Omega : \sup_{k\geq n} \sup_{t\in[0,T]} |\xi_k(t,\omega)| < \varepsilon\}.$$

It is easy to check that $\mathbf{P}(\cup_{n=1}^\infty \Omega_n) = 1$. Then there exists $N = N(\omega)$ such that for all $n > N$

$$\sup_{t\in[0,T]} |\xi_n(t,\omega)| \;=\; \sup_{t\in[0,T]} |\xi_{n\varepsilon}(t,\omega)| \;\leq\; \eta_\varepsilon(\omega) n^{-\gamma},$$

where η_ε is an a.s. finite random variable for every $\varepsilon > 0$. The last inequality holds by the assumptions of the lemma. Define

$$\eta(\omega) \;=\; \sup_{n\geq 1} \sup_{t\in[0,T]} n^\gamma |\xi_n(t,\omega)|.$$

This random variable is a.s. finite. This proves the lemma. □

Lemma 2.4.4 (Gronwall). *Assume that for non-negative increasing continuous processes* y, Q *such that* $Ey(\infty) < \infty$ *and* $Q(\infty) \leq K$ *the following condition holds,*

$$Ey(\tau) \leq E\int_0^\tau y(s)dQ(s) + \varepsilon$$

for any stopping time τ. *Then*

$$Ey(\infty) \leq \varepsilon e^K.$$

2.5 Proof of the Results

2.5.1 Proof of Theorem 2.3.1

We start with rewriting the ordinary and "Stratonovich" equations. Rewrite (2.3.1) in the Itô form

$$dX^i(t) = b_i(t,X(t))dt + \sigma_i^j(t,X(t))dW^j(t) + \frac{1}{2}\sigma^j_{i(\sigma^j)}(t,X(t))dt. \qquad (2.5.1)$$

Writing out the differential for $\sigma_i^j(t,X_n(t))$ and using (2.3.3) we get,

$$\begin{aligned} d\sigma_i^j(t,X_n(t)) &= \frac{\partial}{\partial t}\sigma_i^j(t,X_n(t))dt + b_k(t,X_n(t))\frac{\partial}{\partial x_k}\sigma_i^j(t,X_n(t))dt \\ &\quad + \sigma^j_{i(\sigma^l)}(t,X_n(t))dW_n^l(t). \end{aligned} \qquad (2.5.2)$$

Rewrite (2.3.3) in the form

$$\begin{aligned} dX_n^i(t) &= b_i(t,X_n(t))dt + \sigma_i^j(t,X_n(t))dW^j(t) \\ &\quad + \sigma_i^j(t,X_n(t))d(W_n^j(t) - W^j(t)). \end{aligned}$$

Integrating by parts the last term and applying (2.5.2), we derive

$$\begin{aligned} dX_n^i(t) &= b_i(t,X_n(t))dt \\ &\quad + \sigma_i^j(t,X_n(t))dW^j(t) \\ &\quad + d(\sigma_i^j(t,X_n(t))(W_n^j(t) - W^j(t))) \\ &\quad - (W_n^j(t) - W^j(t))\frac{\partial}{\partial t}\sigma_i^j(t,X_n(t))dt \\ &\quad - (W_n^j(t) - W^j(t))b_k(t,X_n(t))\frac{\partial}{\partial x_k}\sigma_i^j(t,X_n(t))dt \\ &\quad - (W_n^j(t) - W^j(t))\sigma^j_{i(\sigma^l)}(t,X_n(t))dW_n^l(t). \end{aligned} \qquad (2.5.3)$$

Let us fix some positive $\kappa < \alpha$. Define for any integer $R > 0$ stopping times

$$\tau_n^R = \inf\left\{t \ge 0 : |X(t)| + n^\kappa(|W(t) - W_n(t)| + |S_n(t)|) + \frac{||S_n||(t)}{\ln^\delta n} \ge R\right\},$$

$$\pi_n^\varepsilon = \inf\{t \ge 0 : |X_n(t) - X(t)| \ge \varepsilon\},$$

$$\rho_n^{R,\varepsilon} = \tau_n^R \wedge \pi_n^\varepsilon \wedge T.$$

Parameter δ will be chosen later. While using $\rho_n^{R,\varepsilon}$ for simplicity we will omit indices R and ε, and simply write ρ_n. Using (2.5.1), (2.5.3) for any stopping time τ and any $r = 2,3,\ldots$

$$\begin{aligned} &E\sup_{t\le\tau}\left|X^i(t\wedge\rho_n) - X_n^i(t\wedge\rho_n)\right|^r \qquad (2.5.4) \\ &\le c_1\left(I_1 + \sum_j(I_2^j + I_3^j + I_4^j + I_5^j) + \sum_{j,k} I_6^{jk} + \sum_{j,l}(I_7^{jl} + I_8^{jl})\right), \end{aligned}$$

where constant c_1 depends only on r, and

$$\begin{aligned}
I_1 &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(b_i(s,X(s))-b_i(s,X_n(s)))ds\right|^r,\\
I_2^j &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(\sigma_i^j(s,X(s))-\sigma_i^j(s,X_n(s)))dW^j(s)\right|^r,\\
I_3^j &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(\sigma_{i(\sigma^j)}^j(s,X(s))-\sigma_{i(\sigma^j)}^j(s,X_n(s)))ds\right|^r,\\
I_4^j &= E\sup_{t\le\tau}\left|[\sigma_i^j(s,X_n(s))(W_n^j(s)-W^j(s))]_{s=0}^{t\wedge\rho_n}\right|^r,\\
I_5^j &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(W_n^j(s)-W^j(s))\frac{\partial}{\partial t}\sigma_i^j(s,X_n(s))ds\right|^r,\\
I_6^{jk} &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(W_n^j(s)-W^j(s))b_k(s,X_n(s))\frac{\partial}{\partial x_k}\sigma_i^j(s,X_n(s))ds\right|^r,\\
I_7^{jl} &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}(\sigma_{i(\sigma^l)}^j(s,X(s))-\sigma_{i(\sigma^l)}^j(s,X_n(s)))dS_n^{jl}(s)\right|^r,\\
I_8^{jl} &= E\sup_{t\le\tau}\left|\int_0^{t\wedge\rho_n}\sigma_{i(\sigma^l)}^j(s,X(s))dS_n^{jl}(s)\right|^r.
\end{aligned}$$

By Burkholder-Davis-Gundy and Jensen's inequalities as well as by Lipschitz-continuity of b_i, σ_i^j and $\sigma_{i(\sigma^l)}^j$ we get

$$\max_j\{I_1,I_2^j,I_3^j\} \le c_2E\int_0^T\sup_{u\le s}|X(u\wedge\rho_n)-X_n(u\wedge\rho_n)|^r\,d(s\wedge\rho_n),$$

$$\max_{j,l}I_7^{jl} \le c_2(R\ln^\delta n)^{r-1}E\int_0^T\sup_{u\le s}|X(u\wedge\rho_n)-X_n(u\wedge\rho_n)|^r\,d\|S_n^{jl}\|(s\wedge\rho_n),$$

where constant $c_2=c_2(r,T)$ does not depend on n and R. Next, by Assumption 2.3.2 and the definition of the stopping time ρ_n

$$\max_{j,k}\{I_4^j,I_5^j,I_6^{jk}\} \le kE\sup_{t\le\tau\wedge\rho_n}\left|W_n^j(t)-W^j(t)\right|^r \le k_1n^{-\kappa r},$$

where constant $k_1=k_1(r,T,R,\varepsilon)$ does not depend on n.

To estimate I_8^{jl} it suffices to show that $\varphi_n(t) := \sigma_{i(\sigma^l)}^j(t\wedge\rho_n,X(t\wedge\rho_n))$, $\psi_n(t) := S_n^{jl}(t\wedge\rho_n\wedge\tau)$ satisfy the conditions of Lemma 2.4.1 for every n and some p, q, say $p=q=2$. Indeed, using the Itô formula and applying (2.5.1) we derive Itô differential $d\varphi_t^n$, and conclude that by Assumption 2.3.2 and the definition of the stopping time ρ_n functions

$$\begin{aligned} f_n(t) &= \frac{\partial}{\partial t}\sigma^j_{i(\sigma^l)}(t, X(t)) + b_k(t, X(t))\frac{\partial}{\partial x_k}\sigma^j_{i(\sigma_l)}(t, X(t)) \\ &+ \frac{1}{2}\sigma^h_{k(\sigma^h)}(t, X(t))\frac{\partial}{\partial x_k}\sigma^j_{i(\sigma^l)}(t, X(t)) \\ &+ \frac{1}{2}\sigma^h_k(t, X(t))\sigma^h_p(t, X(t))\frac{\partial^2}{\partial x_k \partial x_p}\sigma^j_{i(\sigma^l)}(t, X(t)), \\ g_n(t) &= \sigma^h_k(t, X(t))\frac{\partial}{\partial x_k}\sigma^j_{i(\sigma^l)}(t, X(t)) \end{aligned}$$

are bounded uniformly in $t \in [0, T \wedge \rho_n]$, i.e. assumptions (i)-(iii) of Lemma 2.4.1 hold. Assumption (iv) is satisfied by Assumption 2.3.1. By Lemma 2.4.1

$$\max_{j,l} I_8^{jl} \le k_2 n^{-\kappa r},$$

constant $k_2 = k_2(r, T, R)$ does not depend on n.

Summing up inequalities (2.5.4) by i and using all estimates above we derive

$$Ey_n(\tau) \le cE\int_0^\tau y_n(s)dQ_n(s) + kn^{-\kappa r}, \tag{2.5.5}$$

where constants c, k are independent of n, and

$$y_n(t) = \sup_{u \le t}|X(u \wedge \rho_n) - X_n(u \wedge \rho_n)|^r,$$

$$Q_n(t) = t \wedge \rho_n + (R \ln^\delta n)^{r-1}||S_n||(t \wedge \rho_n)$$

are increasing by t non-negative and continuous processes, and

$$Q_n(t) \le T + (R\ln^\delta n)^r.$$

Using Gronwall lemma (Lemma 2.4.4) for every n we have

$$E\sup_{t \le \rho_n}|X(t) - X_n(t)|^r \le kn^{-\kappa r}\exp\{cT + c(R\ln^\delta n)^r\},$$

which, under condition $\delta r < 1$, implies for every $\beta < \kappa$

$$E\sup_{t \le \rho_n}|X(t) - X_n(t)|^r \le c_\beta n^{-\beta r},$$

where constant $c_\beta = c_\beta(r, T, R)$ does not depend on n.

For any $\gamma < \kappa$ choose β such that $\gamma < \beta < \kappa$, then choose $r > (\beta - \gamma)^{-1}$ and $\delta < 1/r$. By Lemma 2.4.2 almost surely

$$\sup_{t \le \rho_n}|X(t) - X_n(t)| \le \zeta_\gamma n^{-\gamma} \tag{2.5.6}$$

where ζ_γ is a finite random variable.

Below we are getting rid of stopping times τ_n^R and π_n^ε. Let us consider processes

$$A_n(t) \;=\; |X(t)| + n^\kappa(|W(t) - W_n(t)| + |S_n(t)|) \;+\; \frac{||S_n||(t)}{\ln^\delta n},$$

and

$$A(t) \;=\; \sup_n A_n(t).$$

Process $A_n(t)$ is continuous in t for every n, then process $A(t)$ is left-continuous, and by Assumption 2.3.1 almost surely finite for any t. Then

$$\tau^R = \inf\{t \geq 0 : \; A(t) \geq R\}$$

is a stopping time, and inequality $A_n(t) \leq A(t)$ implies

$$\tau_n^R \;\geq\; \tau^R,$$

for every ω and every n. Moreover, almost surely

$$\lim_{R\to\infty} \tau^R \;=\; \infty.$$

By (2.5.6) the random variable

$$\eta_{\gamma,R} \;=\; \sup_n n^\gamma \sup_{t\leq\tau^R\wedge\pi_n^\varepsilon\wedge T} |X(t) - X_n(t)|$$

is almost surely finite, and hence for any $\gamma < \kappa$

$$\sup_{t\leq\tau^R\wedge\pi_n^\varepsilon\wedge T} |X(t) - X_n(t)| \;\leq\; \eta_{\gamma,R} n^{-\gamma}.$$

Consider sets $\Omega_R = \{\omega : \tau^R \geq T\}$. It is obvious that $\mathbf{P}\left(\cup_{R=1}^\infty \Omega_R\right) = 1$. Define $\eta_\gamma = \eta_{\gamma,1}$ for $\omega \in \Omega_1$ and $\eta_\gamma = \eta_{\gamma,R}$ for $\omega \in \Omega_R \backslash (\cup_{k=1}^{R-1}\Omega_R)$, $R \geq 2$. Then η_γ is almost surely finite and

$$\sup_{t\leq\pi_n^\varepsilon\wedge T} |X(t) - X_n(t)| \;\leq\; \eta_\gamma n^{-\gamma}$$

for any $\gamma < \kappa$. Finally, note that because of the arbitrary choice of κ the last inequality holds for any $\gamma < \alpha$. It suffices to apply Lemma 2.4.3. The proof of the theorem is complete.

□

2.5.2 Proof of Theorem 2.3.2

This theorem can be proved in exactly the same way as Theorem 2.3.1. However, under Assumption 2.3.3 the term I_1 in the proof of Theorem 2.3.1 becomes problematic. Our aim now is to get a better expansion for $|X - X_n|$. For simplicity of notations let us denote by $F(t) \cdot G(t)$ the integral $\int_0^t F(s)dG(s)$. For two vectors $u, v \in \mathbb{R}^d$ the scalar product in $\mathbb{R}^d$ we will denote by uv. For simplicity we will also drop parameter t. We get the following expansion. By the Itô formula

$$|X - X_n|^2 = \sum_{k=0}^{4} I_k,$$

where

$$\begin{aligned}
I_0 &= 2\,(X - X_n)(b(t,X) - b(t,X_n)) \cdot t, \\
I_1 &= 2(X - X_n)(\sigma^j(t,X) - \sigma^j(t,X_n)) \cdot W^j, \\
I_2 &= 2(X - X_n)\sigma^j(t,X_n) \cdot (W^j - W^j_n), \\
I_3 &= (X - X_n)\sigma^j_{(\sigma^j)}(t,X) \cdot t, \\
I_4 &= |\,\sigma(t,X)|^2 \cdot t.
\end{aligned}$$

Next, again by the Itô formula

$$I_2 = \sum_{k=0}^{8} I_{2k},$$

where

$$\begin{aligned}
I_{20} &= 2(X - X_n)(W^j - W^j_n)\sigma^j(s,X_n)|_{s=0}^t, \\
I_{21} &= -\,2(W^j - W^j_n)\sigma^j(t,X_n)(b(t,X) - b(t,X_n)) \cdot t, \\
I_{22} &= -\,2(W^j - W^j_n)\sigma^j(t,X_n)(\sigma^l(t,X) - \sigma^l(t,X_n)) \cdot W^l, \\
I_{23} &= -\,2(W^j - W^j_n)\sigma^j(t,X_n)\sigma^l(t,X_n) \cdot (W^l - W^l_n), \\
I_{24} &= -\,(W^j - W^j_n)\sigma^j(t,X_n)\sigma^l_{(\sigma^l)}(t,X) \cdot t, \\
I_{25} &= -\,2(X - X_n)(W^j - W^j_n)\frac{\partial}{\partial t}\sigma^j(t,X_n) \cdot t, \\
I_{26} &= -\,2(X - X_n)(W^j - W^j_n)\frac{\partial}{\partial x_k}\sigma^j(t,X_n)b_k(t,X_n) \cdot t, \\
I_{27} &= -\,2(X - X_n)(W^j - W^j_n)\sigma^j_{(\sigma^l)}(t,X_n) \cdot W^l_n, \\
I_{28} &= -\,2\sigma^j(t,X_n)\sigma^j(t,X) \cdot t;
\end{aligned}$$

$$I_{23} = \sum_{k=0}^{4} I_{23k},$$

where

$$
\begin{aligned}
I_{230} &= -(W^j - W_n^j)(W^l - W_n^l)\sigma^j(s, X_n)\sigma^l(s, X_n)|_{s=0}^t, \\
I_{231} &= 2(W^j - W_n^j)(W^l - W_n^l)\sigma^j(t, X_n)\frac{\partial}{\partial t}\sigma^l(t, X_n) \cdot t, \\
I_{232} &= 2(W^j - W_n^j)(W^l - W_n^l)\sigma^j(t, X_n)\frac{\partial}{\partial x_k}\sigma^l(t, X_n)b_k(t, X_n) \cdot t, \\
I_{233} &= 2(W^j - W_n^j)(W^l - W_n^l)\sigma^j(t, X_n)\sigma^l_{(\sigma^h)}(t, X_n) \cdot W_n^h, \\
I_{234} &= |\,\sigma(t, X_n)|^2 \cdot t.
\end{aligned}
$$

Let us denote

$$
\begin{aligned}
J_1 &= I_{27} + I_3 \;=\; -2(X - X_n)\sigma^j_{(\sigma^l)}(t, X) \cdot S_n^{jl}, \\
J_2 &= I_{233} + I_{24} \;=\; 2(W^j - W_n^j)\sigma^j(t, X_n)\sigma^l_{(\sigma^h)}(t, X) \cdot S_n^{lh}, \\
J_3 &= I_4 + I_{28} + I_{234} \;=\; |\,\sigma(t, X) - \sigma(t, X_n)|^2.
\end{aligned}
$$

We have

$$
J_1 \;=\; \sum_{k=0}^{9} J_{1k},
$$

where

$$
\begin{aligned}
J_{10} &= -2S_n^{jl}(X - X_n)\sigma^j_{(\sigma^l)}(s, X)|_{s=0}^t, \\
J_{11} &= 2S_n^{jl}\sigma^j_{(\sigma^l)}(t, X)(b(t, X) - b(t, X_n)) \cdot t, \\
J_{12} &= 2S_n^{jl}\sigma^j_{(\sigma^l)}(t, X)(\sigma^h(t, X) - \sigma^h(t, X_n)) \cdot W^h, \\
J_{13} &= 2S_n^{jl}\sigma^j_{(\sigma^l)}(t, X)\sigma^h(t, X_n) \cdot (W^h - W_n^h), \\
J_{14} &= S_n^{jl}\sigma^j_{(\sigma^l)}(t, X)\sigma^h_{(\sigma^h)}(t, X) \cdot t, \\
J_{15} &= 2S_n^{jl}(X - X_n)\frac{\partial}{\partial t}\sigma^j_{(\sigma^l)}(t, X) \cdot t, \\
J_{16} &= 2S_n^{jl}(X - X_n)\frac{\partial}{\partial x_p}\sigma^j_{(\sigma^l)}(t, X)b_p(t, X) \cdot t, \\
J_{17} &= 2S_n^{jl}(X - X_n)\frac{\partial}{\partial x_p}\sigma^j_{(\sigma^l)}(t, X)\sigma_p^q(t, X) \cdot W^q, \\
J_{18} &= S_n^{jl}(X - X_n)\frac{\partial}{\partial x_p}\sigma^j_{(\sigma^l)}(t, X)\sigma^q_{p(\sigma^q)}(t, X) \cdot t, \\
J_{19} &= 2S_n^{jl}\frac{\partial}{\partial x_p}\sigma^j_{(\sigma^l)}(t, X)\sigma_p^q(t, X)\sigma^q(t, X) \cdot t;
\end{aligned}
$$

$$
J_2 \;=\; \sum_{k=0}^{9} J_{2k},
$$

where

$$
\begin{aligned}
J_{20} &= 2S_n^{lh}(W^j - W_n^j)\sigma^j(s, X_n)\sigma^l_{(\sigma^h)}(s, X)|_{s=0}^t, \\
J_{21} &= -2S_n^{lh}\sigma^j(t, X_n)\sigma^l_{(\sigma^h)}(t, X) \cdot (W^j - W_n^j), \\
J_{22} &= -2S_n^{lh}(W^j - W_n^j)\sigma^l_{(\sigma^h)}(t, X)\frac{\partial}{\partial t}\sigma^j(t, X_n) \cdot t, \\
J_{23} &= -2S_n^{lh}(W^j - W_n^j)\sigma^l_{(\sigma^h)}(t, X)\frac{\partial}{\partial x_p}\sigma^j(t, X_n)b_p(t, X_n) \cdot t, \\
J_{24} &= -2S_n^{lh}(W^j - W_n^j)\sigma^l_{(\sigma^h)}(t, X)\sigma^j_{(\sigma^q)}(t, X_n) \cdot W_n^q, \\
J_{25} &= -2S_n^{lh}(W^j - W_n^j)\sigma^j(t, X_n)\frac{\partial}{\partial t}\sigma^l_{(\sigma^h)}(t, X) \cdot t, \\
J_{26} &= -2S_n^{lh}(W^j - W_n^j)\sigma^j(t, X_n)\frac{\partial}{\partial x_p}\sigma^l_{(\sigma^h)}(t, X)b_p(t, X) \cdot t, \\
J_{27} &= -2S_n^{lh}(W^j - W_n^j)\sigma^j(t, X_n)\frac{\partial}{\partial x_p}\sigma^l_{(\sigma^h)}(t, X)\sigma_p^q(t, X) \cdot W^q, \\
J_{28} &= -S_n^{lh}(W^j - W_n^j)\sigma^j(t, X_n)\frac{\partial}{\partial x_p}\sigma^l_{(\sigma^h)}(t, X)\sigma^q_{p(\sigma^q)}(t, X) \cdot t, \\
J_{29} &= -2S_n^{lh}\sigma^q(t, X_n)\frac{\partial}{\partial x_p}\sigma^l_{(\sigma^h)}(t, X)\sigma_p^q(t, X) \cdot t.
\end{aligned}
$$

Note that $J_{19} + J_{29} + J_{13} + J_{21} = 0$. Denote

$$
J_4 \;=\; J_{24} + J_{14} \;=\; -2S_n^{jl}\sigma^j_{(\sigma^l)}(t, X)\sigma^h_{(\sigma^q)}(t, X) \cdot S_n^{hq}.
$$

Then

$$
J_4 \;=\; \sum_{k=0}^{5} J_{4k},
$$

where

$$
\begin{aligned}
J_{40} &= -S_n^{jl}\sigma^j_{(\sigma^l)}(s, X)\sigma^h_{(\sigma^q)}(s, X)S_n^{hq}|_{s=0}^t, \\
J_{41} &= 2S_n^{jl}S_n^{hq}\sigma^j_{(\sigma^l)}(t, X)\frac{\partial}{\partial t}\sigma^h_{(\sigma^q)}(t, X) \cdot t, \\
J_{42} &= 2S_n^{jl}S_n^{hq}\sigma^j_{(\sigma^l)}(t, X)\frac{\partial}{\partial x_k}\sigma^h_{(\sigma^q)}(t, X)b_k(t, X) \cdot t, \\
J_{43} &= 2S_n^{jl}S_n^{hq}\sigma^j_{(\sigma^l)}(t, X)\frac{\partial}{\partial x_k}\sigma^h_{(\sigma^q)}(t, X)\sigma_k^p(t, X) \cdot W^p, \\
J_{44} &= S_n^{jl}S_n^{hq}\sigma^j_{(\sigma^l)}(t, X)\frac{\partial}{\partial x_k}\sigma^h_{(\sigma^q)}(t, X)\sigma^p_{k(\sigma^p)}(t, X) \cdot t, \\
J_{45} &= S_n^{jl}S_n^{hq}\frac{\partial}{\partial x_{k_1}}\sigma^j_{(\sigma^l)}(t, X)\frac{\partial}{\partial x_{k_2}}\sigma^h_{(\sigma^q)}(t, X)\sigma^p_{k_1}(t, X)\sigma^p_{k_2}(t, X) \cdot t.
\end{aligned}
$$

Then repeating the proof of Theorem 2.3.1 and denoting

$$
z_n(t) \;=\; \sup_{u \le t} |X(u \wedge \rho_n) - X_n(u \wedge \rho_n)|^{2r}
$$

we can get inequality (2.5.5) with $y_n(t)$ replaced by $z_n(t)$, which leads to the rate $n^{-\gamma/2}$. The first term of the right hand side of this new inequality is formed by $I_0 + J_3$ and the second term is formed by

$$I_1 + \sum_{k=0,1,2,5,6} I_{2k} + \sum_{k=0,1,2} I_{23k} + \sum_{k=0,1,2,5,6,7,8} J_{1k} + \sum_{k=0,2,3,5,6,7,8} J_{2k} + \sum_k J_{4k}.$$

$\square$

Note, that under the Lipschitz-continuity of the problematic terms I_{21}, J_{11} we again get the result $n^{-\gamma}$.

Chapter 3

Stochastic Evolution Equations

3.1 Introduction

In this chapter we consider an abstract form of a stochastic differential equation in infinite dimensional settings often referred to as a stochastic evolution equation. We consider the stochastic differential equation of the form

$$du(t) = (A(t,\omega)u(t) + f(t,\omega))dt + (B^j(t,\omega)u(t) + g^j(t,\omega))dW^j(t) \quad (3.1.1)$$

in a normal triple of spaces $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ (see Definition 3.2.1) with the initial condition

$$u(0) = u_0, \quad (3.1.2)$$

where W is an r-dimensional Wiener process, A, B^j, $j = 1, 2, \dots, r$ are linear operators on $\mathbb{V}$ for every (t,ω), f, g^j are stochastic processes in $\mathbb{H}$ and u_0 is a random variable in $\mathbb{H}$. We recall that here and throughout the paper we use the summation convention with respect to the repeated indices, i.e. in (3.1.1) we sum up the last term with respect to j from 1 to m.

We approximate the Wiener process W with a sequence $\{W_n\}_{n\in\mathbb{N}}$ of continuous processes of finite variation in the supremum norm on the interval $[0,T]$ with the rate $n^{-\kappa}$ for each $\kappa < \alpha$ for fixed $\alpha > 0$, with some additional assumptions on the area process S_n (see Assumption 3.3.1). Simultaneously we approximate operators A, B^j, processes f, g^j, and the initial value u_0 in the same type of topology with the same rate of convergence (see Assumption 3.3.5). We get an approximation sequence of the differential equations

$$du_n(t) = (A_n(t,\omega)u_n(t)+f_n(t,\omega))dt + (B^j_n(t,\omega)u_n(t)+g^j_n(t,\omega))dW^j_n(t), \quad (3.1.3)$$

$n \in \mathbb{N}$, considered in the same triple $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ with the initial conditions

$$u_n(0) = u_{n0}. \quad (3.1.4)$$

The convergence of the sequence of solutions u_n of problems (3.1.3)-(3.1.4) to the solution u of the problem (3.1.1)-(3.1.2) in the supremum norm on $[0, T]$ was shown before (see [6]) under the assumption that the last differential in equation (3.1.1) is interpreted in the Stratonovich sense. We investigate the rate of this convergence. In Theorem 3.3.1 we show that the rate of the convergence $u_n \to u$ is $n^{-\kappa/2}$ for each $\kappa < \alpha$.

The next chapter is devoted to a particular situation of stochastic partial differential equations. There we apply the main result of this chapter to the situation where A is a second order and B^j are a first order differential operators.

3.2 Generalities

Before formulating the result we recall some definitions and fundamental results from the theory of stochastic evolution systems, and introduce some notations. Throughout the chapter for a Banach space, say $\mathbb{U}$, we will denote its norm by $|\cdot|_{\mathbb{U}}$, i.e. we equip the norm sign with the symbol of the space.

3.2.1 Normal Triple

Let $\mathbb{U}$, $\mathbb{V}$ be two Banach spaces. We say that *the space* $\mathbb{U}$ *is normally imbedded into the space* $\mathbb{V}$ (we denote this by $\mathbb{U} \hookrightarrow \mathbb{V}$) if the imbedding is dense (in the topology of the space $\mathbb{V}$ generated by the norm $|\cdot|_{\mathbb{V}}$) and continuous, i.e. there exists a constant N such that $|v|_{\mathbb{V}} \leq N|u|_{\mathbb{U}}$ for any $v \in \mathbb{V}$.

Let $\mathbb{V}$, $\mathbb{V}'$ be two separable Banach spaces, $\mathbb{H}$ be a Hilbert space with the scalar product denoted by $(\cdot\,,\cdot)$.

Definition 3.2.1. *The triple* $(\mathbb{V}, \mathbb{H}, \mathbb{V}')$ *we will call normal and denote by* $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ *if the space* $\mathbb{V}$ *is normally imbedded into the space* $\mathbb{H}$*, which, in return, is normally imbedded into the space* $\mathbb{V}'$*, and for some constant* K

$$|(v, h)| \;\leq\; K|v|_{\mathbb{V}}|h|_{\mathbb{V}'} \tag{3.2.1}$$

for all $v \in \mathbb{V}$, $h \in \mathbb{H}$.

An important example of a normal triple is the Sobolev space $W_2^m(\mathbb{R}^d)$ (the space $\mathbb{V}$), $L_2(\mathbb{R}^d)$ (the space $\mathbb{H}$), and W_2^{-m} (the space $\mathbb{V}'$). Sobolev spaces are considered in more details in Chapter 4.

For any $v' \in \mathbb{V}'$ choose a sequence $\{h_n\}$ from $\mathbb{H}$ such that $|h_n - v'|_{\mathbb{V}'} \to 0$ as $n \to \infty$. By (3.2.1) (v, h_n) converges. Thus for all $v \in \mathbb{V}$ define a bilinear form

$$[v, v'] \;=\; \lim_{n\to\infty} (v, h_n).$$

Clearly $[v, w] = (v, w)$ for $v, w \in \mathbb{H}$. Therefore, below for simplicity we will use the same notations for the bilinear form $[\cdot\,,\cdot]$ and the scalar product in $\mathbb{H}$, and simply denote them by $(\cdot\,,\cdot)$.

It is easy to check that it has the following properties:

(i) it is continuous with respect to both variables, i.e. for all $v \in \mathbb{V}$ and $v' \in \mathbb{V}'$

$$|(v, v')| \leq N|v|_{\mathbb{V}}|v'|_{\mathbb{V}'};$$

(ii) it coincides with the scalar product in $\mathbb{H}$ if $v' \in \mathbb{H}$;

(iii) if $v'(t)$ is an integrable function on the interval $[a, b]$ taking values in $\mathbb{V}'$ then for every $v \in \mathbb{V}$

$$(v, \int_a^b v'(s)ds) = \int_a^b (v, v'(s))ds.$$

The form $(\cdot\,, v')$ defines a linear functional on $\mathbb{V}$ for $v' \in \mathbb{V}'$. Suppose that the equality $(v, v') = 0$ for any $v \in \mathbb{V}$ implies $v' = 0$, and assume that any functional on $\mathbb{V}$ has the form $(\cdot\,, v')$, for some $v' \in \mathbb{V}'$. Then the mapping $\mathcal{J}: v' \to (\cdot\,, v')$ is a one-to-one mapping of $\mathbb{V}'$ onto the space $\mathbb{V}^*$ (here and throughout the paper $\mathbb{V}^*$ denotes the conjugate space for $\mathbb{V}$). Moreover, the following holds.

Proposition 3.2.1. *The mapping $\mathcal{J}$ defines an isometric isomorphism between the spaces $\mathbb{V}'$ and $\mathbb{V}^*$.*

We will say that the scalar product in $\mathbb{H}$ defines the duality between the spaces $\mathbb{V}$ and $\mathbb{V}'$.

3.2.2 Hilbert Scales

Here we recall the definition and basic properties of a Hilbert scale. This topic is covered in more details in [2] (see §7, Chapter II).

Let $\mathbb{V}$, $\mathbb{H}$ be two Hilbert spaces with scalar products $(\cdot\,,\cdot)_{\mathbb{V}}$, $(\cdot\,,\cdot)_{\mathbb{H}}$, and norms $|\cdot|_{\mathbb{V}}$, $|\cdot|_{\mathbb{H}}$, respectively. Let $\mathbb{V}$ be normally imbedded into $\mathbb{H}$, and $|v|_{\mathbb{H}} \leq |v|_{\mathbb{V}}$ for all $v \in \mathbb{V}$. Here we show that this spaces can be connected by a Hilbert scale (see Definition 3.2.2 below), i.e. there exists a Hilbert scale such that one of its elements is the space $\mathbb{V}$ and another is the space $\mathbb{H}$. We also give some properties of the constructed Hilbert scale.

The following is well-known.

Proposition 3.2.2. *There exists a unique self-adjoint positive definite operator Λ defined on the space $\mathbb{V}$ and mapping it onto the space $\mathbb{H}$ such that $|v|_{\mathbb{V}} = |\Lambda v|_{\mathbb{H}}$ for every $v \in \mathbb{V}$.*

The operator Λ is called a *generating* operator for the pair $(\mathbb{H}, \mathbb{V})$.

Making use of the spectral decomposition of the identity E_λ corresponding to the operator Λ we define powers of this operator by the formula

$$\Lambda^\alpha v = \int_0^\infty \lambda^\alpha dE_\lambda v, \qquad \alpha \in \mathbb{R}$$

(see §3, Chapter II, [2]).

For every $\alpha \geq 0$ the domain of operator Λ^α we denote by $\mathbb{H}_\alpha$. For $\alpha < 0$ we define $\mathbb{H}_\alpha$ as the completion of the space $\mathbb{H}$ with respect to the norm $|\cdot|_{\mathbb{H}_\alpha} = |\Lambda^\alpha \cdot|_{\mathbb{H}}$. It is known that $\mathbb{H}_\alpha$ defined in this way is a Hilbert space for each $\alpha \in \mathbb{R}$ with respect to the scalar product

$$(\cdot\,,\cdot)_{\mathbb{H}_\alpha} = (\Lambda^\alpha \cdot\,, \Lambda^\alpha \cdot).$$

Definition 3.2.2. *A system of Hilbert spaces $\{\mathbb{H}_\alpha\}_{\alpha\in\mathbb{R}}$ is called a Hilbert scale if for any three real α, β, γ such that $\alpha < \beta < \gamma$ the space $\mathbb{H}_\gamma$ is normally imbedded into $\mathbb{H}_\beta$ and in turn $\mathbb{H}_\beta$ is normally imbedded into $\mathbb{H}_\alpha$, and for every $v \in \mathbb{H}_\gamma$*

$$|v|_{\mathbb{H}_\beta} \leq |v|_{\mathbb{H}_\alpha}^{(\gamma-\beta)/(\gamma-\alpha)} |v|_{\mathbb{H}_\gamma}^{(\beta-\alpha)/(\gamma-\alpha)}.$$

It turns out that by the representation for the operators Λ^α the system $\{\mathbb{H}_\alpha\}_{\alpha\in\mathbb{R}}$ is a Hilbert scale possessing the following properties,

(i) it is uniquely defined;

(ii) $\mathbb{H}_\infty = \bigcap_{\beta\in\mathbb{R}} \mathbb{H}_\beta$ is dense in the space $\mathbb{H}_\alpha$ for all $\alpha \in \mathbb{R}$ in the norm $|\cdot|_{\mathbb{H}_\alpha}$;

(iii) the spaces $\mathbb{H}_\alpha$ and $\mathbb{H}_{-\alpha}$ are mutually conjugate with respect to the scalar product in $\mathbb{H}_0$.

The system $\{\mathbb{H}_\alpha\}_{\alpha\in\mathbb{R}}$ is called *the Hilbert scale connecting spaces* $\mathbb{V}$ *and* $\mathbb{H}$.

Let us choose real α, β such that $\alpha < \beta$ and define $\gamma = 2\beta - \alpha$. We consider the triple of spaces $(\mathbb{H}_\alpha, \mathbb{H}_\beta, \mathbb{H}_\gamma)$. For any $v \in \mathbb{H}_\alpha$, $h \in \mathbb{H}_\beta$

$$\begin{aligned} |(v,h)_{\mathbb{H}_\beta}| &= |(\Lambda^\beta v, \Lambda^\beta h)_{\mathbb{H}_0}| = |(\Lambda^{\beta-\alpha}\Lambda^\alpha v, \Lambda^\beta h)_{\mathbb{H}_0}| \\ &= |(\Lambda^\alpha v, \Lambda^\gamma h)_{\mathbb{H}_0}| \leq |v|_{\mathbb{H}_\alpha} |h|_{\mathbb{H}_\gamma}, \end{aligned}$$

and hence we have the following.

Proposition 3.2.3. *The system $(\mathbb{H}_\alpha, \mathbb{H}_\beta, \mathbb{H}_\gamma)$ is a normal triple.*

In view of the last assertion, by the properties of normal triples (see Proposition 3.2.1), the scalar product in $\mathbb{H}_\beta$ defines the duality mapping between $\mathbb{H}_\alpha$ and $\mathbb{H}_\gamma$.

3.2.3 Itô formula

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space equipped with a complete right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$, i.e. an expanding system of σ-algebras imbedded in $\mathcal{F}$. Let us fix a normal triple $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$. We will use the following fundamental result.

Theorem 3.2.4 (Itô formula). *Let $\tilde{v}_i$ be a $\mathbb{V}'$-valued stochastic process for $i = 1, 2$ such that almost surely*

$$\tilde{v}_i(t) \;=\; v_{i0} \;+\; \int_0^t v_i'(s)dN(s) \;+\; m_i(t)$$

for all $t \in [0,T]$, where v_{i0} is an $\mathbb{H}$-valued $\mathcal{F}_0$-measurable random variable, N is an $\mathcal{F}_t$-adapted continuous stochastic process of bounded variation, m_i is a continuous $\mathcal{F}_t$-adapted local martingale for each $i = 1, 2$. Assume there exists a $\mathbb{V}$-valued $\mathcal{F}_t$-adapted process v_i such that

$$\tilde{v}_i(t) \;=\; v_i(t)$$

for $dN \times d\mathbf{P}$-almost every $(t, \omega) \in [0,T] \times \Omega$. Assume, moreover, that almost surely

$$\int_0^T (|v(s)|_{\mathbb{V}} + |\tilde{v}(s)|_{\mathbb{V}'} + |v(s)|_{\mathbb{V}}|\tilde{v}(s)|_{\mathbb{V}'})dN(s) \;<\; \infty.$$

Then there exists a set $\Omega' \subset \Omega$ such that $\mathbf{P}(\Omega') = 1$, $\tilde{v}_i(t,\omega) \in \mathbb{H}$ for all $t \in [0,T]$, $\omega \in \Omega'$, and for all $\omega \in \Omega'$

$$\begin{aligned}
(\tilde{v}_1(t), \tilde{v}_2(t)) \;=\;& (v_1(0), v_2(0)) \\
&+ \int_0^t (v_1(s), v_2'(s))dN(s) \;+\; \int_0^t (v_1'(s), v_2(s))dN(s) \\
&+ \int_0^t (\tilde{v}_1(s), dm_2(s)) \;+\; \int_0^t (\tilde{v}_2(s), dm_1(s)) \\
&+ \langle m_1, m_2\rangle(t).
\end{aligned}$$

This theorem follows directly from Theorem 3.2. (Itô formula for the square of the norm), [9] by making use of the formula $(\tilde{v}_1, \tilde{v}_2) = \frac{1}{4}(|\tilde{v}_1 + \tilde{v}_2|^2 - |\tilde{v}_1 - \tilde{v}_2|^2)$. Using Theorem 3.2.4 we say that we apply the Itô formula to the scalar product (v_1, v_2) in the triple $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$.

3.2.4 Stochastic Evolution Equations

In this chapter we consider the stochastic differential equation of the form

$$v(t) \;=\; v_0 \;+\; \int_0^t \mathbb{A}(s, v(s))ds \;+\; \int_0^t \mathbb{B}^j(s, v(s))dW^j(s) \tag{3.2.2}$$

in a normal triple $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ on a given complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a complete right-continuous filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, $(T > 0)$. The equation is considered on the interval $[0, T]$. Above W is an r-dimensional Wiener process, v_0 is an $\mathcal{F}_0$-measurable random variable with values in $\mathbb{H}$, and $\mathbb{A}$, $\mathbb{B}^j$, $j = 1, \ldots, r$, are random fields taking values in $\mathbb{V}'$ and $\mathbb{H}$, respectively, such that for each $v \in \mathbb{V}$ the functions $\mathbb{A}(t, \omega)v$, $\mathbb{B}(t, \omega)v$ are measurable in (t, ω) (relative to the measure $dt \times d\mathbf{P}$) and $\mathcal{F}_t$-consistent, i.e. they are $\mathcal{F}_t$-measurable in ω for each $v \in \mathbb{V}$, $t \in [0, T]$.

Definition 3.2.3. *We will call an* $\mathbb{H}$*-solution of the equation (3.2.2) in the triple* $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ *on the interval* $[0, T]$ *an* $\mathbb{H}$*-valued continuous* $\mathcal{F}_t$*-adapted stochastic process* v *defined on* $[0, T]$ *if*

- *(i)* $v(t, \omega) \in \mathbb{V}$ *for* $dt \times \mathbf{P}$ *almost every* (t, ω)*;*
- *(ii) there exists a set* $\Omega' \subset \Omega$ *with* $\mathbf{P}(\Omega') = 1$ *such that (3.2.2) holds for every* $\omega \in \Omega'$ *and* $t \in [0, T]$*, where the equality is understood as the equality of elements of* $\mathbb{V}'$*;*
- *(iii)* $\int_0^T |v(t)|_{\mathbb{V}}^2 dt < \infty$ *almost surely.*

Although this is not the topic of the present work, we next give assertions on the existence and uniqueness of the $\mathbb{H}$-solution for equation (3.2.2). This topic is described in more details in [18], [22] (Chapter 3). The existence result is formulated under the following assumptions.

Assumption 3.2.1.

- *(i) Semicontinuity of* $\mathbb{A}$*: the function* $v\mathbb{A}(t, v_1 + \lambda v_2)$ *is continuous in* λ *on* $\mathbb{R}$ *for all* $(t, \omega) \in [0, T] \times \Omega$*;*
- *(ii) Monotonicity of* $(\mathbb{A}, \mathbb{B})$*: there exists a constant* K *such that*

$$2(v_1 - v_2, \mathbb{A}(t, v_1) - \mathbb{A}(t, v_2)) \ + \ |\mathbb{B}^j(t, v_1) - \mathbb{B}^j(t, v_2)|_{\mathbb{H}}^2 \ \leq \ K|v_1 - v_2|_{\mathbb{H}}^2$$

for all $v_1, v_2 \in \mathbb{V}$*,* $(t, \omega) \in [0, T] \times \Omega$*;*

- *(iii) Coercivity of* $(\mathbb{A}, \mathbb{B})$*: there exist constants* K*,* $\alpha > 0$ *such that*

$$2(v, \mathbb{A}(t, v)) \ + \ |\mathbb{B}(t, v)|_{\mathbb{H}}^2 \ + \ \alpha|v|_{\mathbb{V}}^2 \ \leq \ K|v|_{\mathbb{H}}^2$$

for all $v \in \mathbb{V}$*,* $(t, \omega) \in [0, T] \times \Omega$*;*

(iv) *Boundedness of the growth of* $\mathbb{A}$: *there exists a constant* K *such that*

$$|\mathbb{A}(t,v)|_{\mathbb{V}'} \le K|v|_{\mathbb{V}}$$

for all $v \in \mathbb{V}$, $(t,\omega) \in [0,T] \times \Omega$;

(v)

$$E|v_0|^2_{\mathbb{H}} < \infty.$$

The following two theorems are proved in [18].

Theorem 3.2.5. *Under Assumption 3.2.1 there exists an* $\mathbb{H}$*-solution of the equation (3.2.2) in the triple* $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ *on the interval* $[0,T]$.

The next theorem implies, in particular, the uniqueness assertion for the solution of equation (3.2.2).

Theorem 3.2.6. *Let* v_n, $n = 0, 1, \dots,$ *be* $\mathbb{H}$*-solutions of equation (3.2.2) with the initial condition* v_{n0} *in place of* v_0, *where* $E|v_{n0}|^2_{\mathbb{H}} < \infty$ *and* $E|u_{00} - u_{n0}|^2_{\mathbb{H}} \to 0$ *as* $n \to \infty$. *Then for any* $\varepsilon > 0$

$$\lim_{n\to\infty} (\sup_{t\le T} E|v_0(t) - v_n(t)|^2_{\mathbb{H}} + \mathbf{P}\{\sup_{t\le T} |v_0(t) - v_n(t)|_{\mathbb{H}} \ge \varepsilon\}) = 0.$$

3.2.5 Other Notations

We will also need the following concepts. Let $\mathbb{U}$, $\mathbb{V}$ be two separable Banach spaces. We will denote by $\mathbb{L}(\mathbb{U},\mathbb{V})$ the vector space of bounded linear operators mapping $\mathbb{U}$ into $\mathbb{V}$. Stochastic process X will be called *measurable in* $\mathbb{L}(\mathbb{U},\mathbb{V})$ if $X(t,\omega)$ belongs to $\mathbb{L}(\mathbb{U},\mathbb{V})$ for every $(t,\omega) \in [0,\infty) \times \Omega$, and Xu is a measurable process in $\mathbb{V}$ for every $u \in \mathbb{U}$. Stochastic process Y will be called *continuous in* $\mathbb{L}(\mathbb{U},\mathbb{V})$ if $Y(t)u$ is a continuous process in the space $\mathbb{V}$ for all $u \in \mathbb{U}$.

Similarly to the notion of O introduced in Chapter 2 we define the following.

(i) For a sequence of stochastic processes $\{f_n\}_{n\in\mathbb{N}}$ with values in $\mathbb{V}$ and a numerical sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ we write $f_n = O(\alpha(n))$ *in* $\mathbb{V}$ if $|f_n(t)|_{\mathbb{V}} \le \zeta\alpha(n)$ for every $t \in [0,T]$, for some almost surely finite random variable ζ which does not depend on t, n.

(ii) For a sequence of stochastic processes $\{X_n\}_{n\in\mathbb{N}}$ with values in $\mathbb{L}(\mathbb{U},\mathbb{V})$ and a numerical sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ we write $X_n = O(\alpha(n))$ *in* $\mathbb{L}(\mathbb{U},\mathbb{V})$ if $X_n u = O(\alpha(n))$ in $\mathbb{V}$ for all $u \in \mathbb{U}$.

Let the space $\mathbb{U}$ be normally imbedded into the space $\mathbb{V}$, $\mathbb{U} \hookrightarrow \mathbb{V}$. A random process $v = \{v(t)\}_{t\in[0,T]}$ is said to belong to the class $\mathcal{C}([0,T];\mathbb{V}) \cap \mathcal{L}_2([0,T];\mathbb{U})$ if v is an $\mathcal{F}_t$-adapted continuous process in $\mathbb{V}$ and there exists its $\mathbb{U}$-version u (i.e. process u with values in $\mathbb{U}$ such that almost surely $v(t) = u(t)$ for every $t \in [0,T]$), for which $\int_0^T |u(t)|_{\mathbb{U}}^2 dt < \infty$.

We call a stochastic process $v = \{v(t)\}_{t\geq 0}$ taking values in a separable Banach space $\mathbb{U}$ a *process of bounded variation* if for every $t \geq 0$ the total variation of v over the interval $[0,t]$,

$$||v||(t) \;=\; \sup \sum_{k=1}^{n} |v(t_k) - v(t_{k-1})|_{\mathbb{U}}$$

is almost surely finite. Above the least upper bound is taking over all (finite) partitions $0 = t_0 < t_1 < \cdots < t_n = t$ of the interval $[0,t]$.

3.3 The Main Result

Given a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a complete right-continuous filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, $T > 0$, let W be an r-dimensional Wiener process and $\{W_n\}_{n\in\mathbb{N}}$ a sequence of r-dimensional continuous processes of bounded variation both defined on the interval $[0,T]$. Let us fix some positive number α. Suppose that the following assumption holds.

Assumption 3.3.1. *For any positive $\kappa < \alpha$ and every $\delta > 0$*

$$\begin{aligned}
&(1) \qquad W - W_n \;=\; O(n^{-\kappa}),\\
&(2) \qquad S_n \;=\; O(n^{-\kappa}),\\
&(3) \qquad ||S_n|| \;=\; O(\ln^{\delta} n),
\end{aligned}$$

where S_n is an $r \times r$-dimensional process defined as follows,

$$S_n^{jl}(t) \;=\; \int_0^t (W^j(s) - W_n^j(s))dW_n^l(s) \;-\; \frac{1}{2}\delta_{jl}\, t,$$

where δ_{jl} is the Kronecker's symbol which assumes 1 if $j = l$, and 0 otherwise.

Let $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$ be a normal triple. The scalar product in $\mathbb{H}_0$ we will denote as $(\cdot\,,\cdot)$.

Let A, B, f and g be well-measurable stochastic processes on $[0,T]$ assuming their values in $\mathbb{L}(\mathbb{H}_1, \mathbb{H}_{-1})$, $\mathbb{L}(\mathbb{H}_1, \mathbb{H}_0^m)$, $\mathbb{H}_{-1}$ and $\mathbb{H}_0^m$, respectively (here $\mathbb{H}_0^m$ denotes the product space $\mathbb{H}_0 \times \cdots \times \mathbb{H}_0$, m times), and let A_n, B_n, f_n and g_n, $n \in \mathbb{N}$, be their approximation sequences. Suppose the following holds.

Assumption 3.3.2. *For every $j = 1, 2, \ldots, r$ there exist measurable stochastic processes $(B^{j(0)}, \ldots, B^{j(m)})$, $(B_n^{j(0)}, \ldots, B_n^{j(m)})$, $n \in \mathbb{N}$, taking values in $\mathbb{L}(\mathbb{H}_1, \mathbb{H}_0^{m+1})$ and processes $(g^{j(0)}, \ldots, g^{j(m)})$, $(g_n^{j(0)}, \ldots, g_n^{j(m)})$, $n \in \mathbb{N}$, taking values in $\mathbb{H}_0^{m+1}$, such that the processes $(v, B^j(t)w)$, $(v, B_n^j(t)w)$, $(v, g^j(t))$ and $(v, g_n^j(t))$ have for every $v, w \in \mathbb{H}_1$ stochastic differentials*

$$\begin{aligned} d(v, B^j(t)w) &= (v, B^{j(0)}w)dt + (v, B^{j(l)}w)dW^l(t), \\ d(v, B_n^j(t)w) &= (v, B_n^{j(0)}w)dt + (v, B_n^{j(l)}w)dW_n^l(t), \end{aligned}$$

$$\begin{aligned} d(v, g^j(t)) &= (v, g^{j(0)})dt + (v, g^{j(l)})dW^l(t), \\ d(v, g_n^j(t)) &= (v, g_n^{j(0)})dt + (v, g_n^{j(l)})dW_n^l(t). \end{aligned}$$

Moreover, for every $j, k = 1, 2, \ldots, r$ there exist measurable stochastic processes $(B^{j(k0)}, \ldots, B^{j(km)})$ taking values in $\mathbb{L}(\mathbb{H}_1, \mathbb{H}_0^{m+1})$ and processes $(g^{j(k0)}, \ldots, g^{j(km)})$ taking values in $\mathbb{H}_0^{m+1}$, such that the processes $(v, B^{j(k)}(t)w)$, $(v, g^{j(k)}(t))$ have for every $v, w \in \mathbb{H}_1$ stochastic differentials

$$d(v, B^{j(k)}(t)w) = (v, B^{j(k0)}w)dt + (v, B^{j(kl)}w)dW^l(t),$$

$$d(v, g^{j(k)}(t)) = (v, g^{j(k0)})dt + (v, g^{j(kl)})dW^l(t).$$

We stress that $k \neq 0$ above.

Let $\mathbb{H}_{\pm 2}, \mathbb{H}_{\pm 3}, \mathbb{H}_4, \mathbb{H}_5, \mathbb{H}_6$ be separable Banach spaces such that $\mathbb{H}_k$ is normally imbedded into $\mathbb{H}_{k-1}$, $k = -1, 0, \ldots, 6$, and $\mathbb{H}_k \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-k}$ form a normal triple for every $k = 1, 2, 3$. We have

$$\mathbb{H}_6 \hookrightarrow \mathbb{H}_5 \hookrightarrow \mathbb{H}_4 \hookrightarrow \mathbb{H}_3 \hookrightarrow \mathbb{H}_2 \hookrightarrow \mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1} \hookrightarrow \mathbb{H}_{-2} \hookrightarrow \mathbb{H}_{-3}.$$

We use the following notations. The norm in $\mathbb{H}_\beta$ we denote by $|\cdot|_\beta$. The scalar product in $\mathbb{H}_0$ is denoted by $(\cdot, \cdot)$. For a linear operator X mapping $\mathbb{H}_\beta$ into $\mathbb{H}_\gamma$ its operator norm is denoted by $|X|_{\beta,\gamma}$. If operator X is defined on a subspace that is dense in $\mathbb{H}_\beta$ then $|X|_{\beta,\gamma}$ denotes the norm of X extended by continuity to the whole space $\mathbb{H}_\beta$.

We consider "Stratonovich" stochastic differential equation

$$u(t) = \xi + \int_0^t (A(s)u(s) + f(s))ds + \int_0^t (B^j(s)u(s) + g^j(s)) \circ dW^j(s) \quad (3.3.1)$$

in the triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$, where ξ is an $\mathcal{F}_0$-measurable random variable with values in $\mathbb{H}_2$. Here we used the short notation for the infinite-dimensional analogue of the finite-dimensional Stratonovich differential,

$$\begin{aligned}(B^j(t)u(t)+g^j(t))\circ dW^j(t) &= (B^j(t)u(t)+g^j(t))dW^j(t)\\ &+\frac{1}{2}B^j(t)(B^j(t)u(t)+g^j(t))dt\\ &+\frac{1}{2}(B^{j(j)}(t)u(t)+g^{j(j)}(t))dt\end{aligned}$$

for every fixed j. For $n \in \mathbb{N}$ we consider differential equations

$$u_n(t) = \xi_n + \int_0^t (A_n(s)u_n(s)+f_n(s))ds + \int_0^t (B_n^j(s)u_n(s)+g_n^j(s))dW_n^j(s) \tag{3.3.2}$$

in the same triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$, where ξ_n is an $\mathcal{F}_0$-measurable random variable with values in $\mathbb{H}_1$.

We assume the following.

Assumption 3.3.3. *There exists a finite random variable ζ such that*

$$\begin{aligned}
|A(t)|_{k,k-2} &\le \zeta, & k&=0,1,\dots,6, & |f(t)|_4 &\le \zeta,\\
|B^j(t)|_{k,k-1} &\le \zeta, & k&=-2,-1,\dots,6, & |g^j(t)|_5 &\le \zeta,\\
|B^{j(l)}(t)|_{k,k-1} &\le \zeta, & k&=-2,-1,\dots,5, & |g^{j(l)}(t)|_4 &\le \zeta,\\
|B^{j(kl)}(t)|_{3,2} &\le \zeta, & & & |g^{j(kl)}(t)|_2 &\le \zeta,
\end{aligned}$$

and for every $n \in \mathbb{N}$

$$\begin{aligned}
|A_n(t)|_{k,k-2} &\le \zeta, & k&=0,1,2, & |f_n(t)|_0 &\le \zeta,\\
|B_n^j(t)|_{k,k-1} &\le \zeta, & k&=-1,0,1, & |g_n^j(t)|_1 &\le \zeta,\\
|B_n^{j(l)}(t)|_{k,k-1} &\le \zeta, & k&=-1,0, & |g_n^{j(l)}(t)|_0 &\le \zeta,
\end{aligned}$$

for every $j,k=1,2,\dots,r$, $l=0,1,\dots,r$ and $(t,\omega)\in[0,T]\times\Omega$.

Assumption 3.3.4. *There exists an almost surely finite random variable ζ such that for some positive λ for every $j,l=1,2,\dots,r$ and $n\in\mathbb{N}$*

(i) for every $v \in \mathbb{H}_1$

$$\begin{aligned}
(v,A_nv) + \lambda|v|_1^2 &\le \zeta|v|_0^2,\\
|(B_n^jv,B_n^lv) + (v,B_n^jB_n^lv)| &\le \zeta|v|_0^2,\\
|(v,B_n^jv)| &\le \zeta|v|_0^2,\\
|(v,B_n^{j(l)}v)| &\le \zeta|v|_0^2,
\end{aligned}$$

(ii) for every $v \in \mathbb{H}_2$

$$|(A_n v, B_n^j v) + (v, B_n^j A_n v)| \leq \zeta |v|_1^2.$$

The first inequality in the assumption above is the coercivity of the operator A_n which we need for the existence of the solution for the equation (3.3.1) in the triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$ (compare with Assumption 3.2.1). The other inequalities look less usual, however, are satisfied in Sobolev spaces by B_n^j defined as the first order and A_n defined as the second order differential operators. This situation is the subject of Chapter 4.

Assumption 3.3.5.

(i) For every $\kappa < \alpha$ for all $j, l = 1, 2, \ldots, r$

$$\begin{aligned} A - A_n &= O(n^{-\kappa}) \text{ in } \mathbb{L}(\mathbb{H}_1, \mathbb{H}_{-1}), \\ B^j - B_n^j &= O(n^{-\kappa}) \text{ in } \mathbb{L}(\mathbb{H}_2, \mathbb{H}_1) \text{ and } \mathbb{L}(\mathbb{H}_1, \mathbb{H}_0), \\ B^{j(l)} - B_n^{j(l)} &= O(n^{-\kappa}) \text{ in } \mathbb{L}(\mathbb{H}_1, \mathbb{H}_0); \end{aligned}$$

(ii) For every $\kappa < \alpha$ for all $j, l = 1, 2, \ldots, r$

$$\begin{aligned} f - f_n &= O(n^{-\kappa}) \text{ in } \mathbb{H}_0, \\ g^j - g_n^j &= O(n^{-\kappa}) \text{ in } \mathbb{H}_1, \\ g^{j(l)} - g_n^{j(l)} &= O(n^{-\kappa}) \text{ in } \mathbb{H}_0; \end{aligned}$$

(iii) for every $\kappa < \alpha$

$$\xi - \xi_n = O(n^{-\kappa}) \text{ in } \mathbb{H}_0.$$

Note that in two assumptions above l does not assume value 0.

Assumption 3.3.6. *On the interval $[0, T]$ there exist solutions u, u_n of the equations (3.3.1), (3.3.2), respectively, such that u is from the class $\mathcal{C}([0,T]; \mathbb{H}_5) \cap \mathcal{L}_2([0,T]; \mathbb{H}_6)$, u_n is from the class $\mathcal{C}([0,T]; \mathbb{H}_1) \cap \mathcal{L}_2([0,T]; \mathbb{H}_2)$.*

Theorem 3.3.1. *Under Assumptions 3.3.1-3.3.6 the sequence of solutions u_n of differential equations (3.3.2) converges almost surely to the solution u of "Stratonovich" stochastic differential equation (3.3.1). Moreover, for every $\kappa < \alpha$*

$$|u - u_n|_0^2 = O(n^{-\kappa}),$$

and

$$\int_0^T |u(s) - u_n(s)|_1^2 ds = O(n^{-\kappa}).$$

3.4 Auxiliary Results

In this section we give some auxiliary statements. They are to be used to prove the main result of the chapter. The following lemma is similar to Lemma 2.4.2. However, processes are considered in an arbitrary Banach space.

Lemma 3.4.1. *Let β, γ be two positive numbers, $\gamma < \beta$. Assume that for a sequence of stochastic processes $\{B_n(t)\}_{n\in\mathbb{N}}$, $t \in [0,T]$, with values in a separable Banach space $\mathbb{V}$, for every n and some $r > (\beta-\gamma)^{-1}$ the following condition holds,*

$$(E \sup_{t\leq T} |B_n(t)|_{\mathbb{V}}^r)^{1/r} \;\leq\; c_\beta n^{-\beta},$$

where c_β may depend on r but not on n. Then

$$B_n \;=\; O(n^{-\gamma}) \text{ in } \mathbb{V}.$$

Proof. We simply set $\bar{B}(t) = |B(t)|_{\mathbb{V}}$ and apply Lemma 2.4.2 for $\bar{B}$. □

The following statement is a corollary of the Burkholder-Davis-Gundy estimation.

Lemma 3.4.2. *Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a sequence of stochastic processes with values in a Banach space $\mathbb{V}$. Let α be a positive number. Suppose that for every positive $\kappa < \alpha$*

$$\varphi_n(t) \;=\; O(n^{-\kappa}) \text{ in } \mathbb{V}. \tag{3.4.1}$$

Let W be a Wiener process, and assume the existence for every n of the stochastic integral of φ_n with respect to W. Then for every positive $\gamma < \alpha$

$$\int_0^t \varphi_n(s)dW(s) \;=\; O(n^{-\gamma}) \text{ in } \mathbb{V}. \tag{3.4.2}$$

Proof. Let us fix any κ. Denote $G_n(t) = n^\kappa |\varphi_n(t)|_{\mathbb{V}}$. Let us define for any positive integer R a sequence of stopping times

$$\tau_R^n \;=\; \inf\{t \geq 0,\ G_n(t) \geq R\}.$$

By Burkholder inequality for every number $r = 2, 3, \ldots$

$$E \sup_{t\leq\tau_R^n\wedge T} \left|\int_0^t \varphi_n(s)dW(s)\right|_{\mathbb{V}}^r \;\leq\; c_r E \sup_{t\leq\tau_R^n\wedge T} |\varphi_n(t)|_{\mathbb{V}}^r \;\leq\; Cn^{-\kappa r},$$

where constant $C = C(r,R)$ does not depend on n. Then by Lemma 3.4.1, assigning $B^n(t) = \int_0^{t\wedge\tau_R^n} \varphi_n(s)dW(s)$, for every positive $\gamma < \kappa$

$$\sup_{t\leq\tau_R^n\wedge T} \left|\int_0^t \varphi_n(s)dW(s)\right|_{\mathbb{V}} \;\leq\; \eta_{\gamma,R} n^{-\gamma}, \tag{3.4.3}$$

where $\eta_{\gamma,R}$ is an a.s. finite random variable. Denote $G(t) = \sup_n G_n(t)$ and define stopping time

$$\tau_R \;=\; \inf\{t \geq 0,\; G(t) \geq R\}$$

Inequality $G_n(t) \leq G(t)$ implies relation

$$\tau_R \;\leq\; \tau_R^n$$

and, therefore, by (3.4.3), for any $\gamma < \kappa$

$$\sup_{t \leq \tau_R \wedge T} \left| \int_0^t \varphi_n(s) dW(s) \right|_{\mathbb{V}} \;\leq\; \eta_{\gamma,R} n^{-\gamma}.$$

By assumption (3.4.1) of the lemma process $G(t)$ is bounded above by an almost surely finite random variable. This implies almost surely

$$\lim_{R\to\infty} \tau_R \;=\; \infty.$$

Introduce sets $\Omega_R = \{\omega,\; \tau_R \geq T\}$. Obviously, almost every ω hits some Ω_R. So that, $\mathbf{P}(\cup_{R=1}^{\infty}\Omega_R) \;=\; 1$. Define $\eta_\gamma \;=\; \eta_{\gamma,R}$ for $\omega \in \Omega_R \backslash (\cup_{k=1}^{R-1}\Omega_k)$, $R \;=\; 1, 2, \ldots$. For other values of ω define $\eta_\gamma = \infty$. Then η_γ is an almost surely finite random variable, and for any $\gamma < \kappa$ inequality (3.4.2) holds.

Finally, note that because κ can be chosen arbitrarily inequality (3.4.2) holds for any $\gamma < \alpha$. □

The following two lemmas are technical tools for the estimation of the terms of a special form. Let y be a stochastic process with values in a Hilbert space $\mathbb{V}$. Assume that process y has a stochastic differential

$$dy(t) \;=\; a(t)dt \;+\; b^j(t)dW^j(t),$$

where W is an r-dimensional Wiener process. Suppose that processes y, a, b^j, $j = 1, 2, \ldots, r$ satisfy the following conditions,

$$y, b^j \;=\; O(1) \text{ in } \mathbb{V}, \qquad \int_0^T |a(t)|_{\mathbb{V}}^2 dt \;=\; O(1).$$

Let α be a positive number, and $\{\psi_n\}_{n\in\mathbb{N}}$ a sequence of real valued stochastic processes, such that for every $\kappa < \alpha$

$$\psi_n(t) \;=\; O(n^{-\kappa}).$$

Lemma 3.4.3. *Under the assumptions above for every $\kappa < \alpha$*

$$\int_0^t y(s) d\psi_n(s) \;=\; O(n^{-\kappa})$$

providing that this integral exists.

Proof. We rewrite the assumptions of the lemma in the form

$$
\begin{aligned}
\sup_{t\le T} |y(t)|_{\mathbb{V}} &\le \zeta, \\
\int_0^T |a(t)|^2_{\mathbb{V}} dt &\le \zeta, \\
\sup_{t\le T} |b^j(t)|_{\mathbb{V}} &\le \zeta,
\end{aligned}
$$

for an a.s. finite random variable ζ,

$$
\sup_{t\le T} |\psi_n(t)| \le \xi_\kappa n^{-\kappa}
$$

for every positive $\kappa < \alpha$ and some a.s. finite (for every κ) random variables ξ_κ.

Integrating by parts we get

$$
\int_0^t y(s) d\psi_n(s) = \sum_{k=1}^3 \theta_n^k(t),
$$

where

$$
\begin{aligned}
\theta_n^1(t) &= y(t)\psi_n(t) - y(0)\psi_n(0), \\
\theta_n^2(t) &= -\int_0^t a(s)\psi_n(s)ds, \\
\theta_n^3(t) &= -\int_0^t b^j(s)\psi_n(s)dW^j(s).
\end{aligned}
$$

Below we estimate the terms θ_n^k.

$$
|\theta_n^1(t)|_{\mathbb{V}} \le 2\sup_t |y(t)|_{\mathbb{V}} \sup_t |\psi_n(t)| \le \zeta\xi_\kappa n^{-\kappa}.
$$

$$
|\theta_n^2(t)|_{\mathbb{V}} \le \left(\int_0^t |a(s)|^2_{\mathbb{V}} ds\right)^{1/2} \left(\int_0^t |\psi_n(s)|^2 ds\right)^{1/2} \le \sqrt{\zeta T}\xi_\kappa n^{-\kappa}.
$$

Assign $\varphi_n(t) = b^j(t)\psi_n(t)$. By assumptions of the lemma $\sup_t |\varphi_n(t)|_{\mathbb{V}} \le \zeta\xi_\kappa n^{-\kappa}$. Then by Lemma 3.4.2

$$
|\theta_n^3(t)|_{\mathbb{V}} \le \eta_\kappa n^{-\kappa},
$$

where random variable η_κ is almost surely finite. It suffices to notice that the random variable $\zeta\xi_\kappa + \sqrt{\zeta T}\xi_\kappa + \eta_\kappa$ is a.s. finite. $\square$

Let $\mathbb{H}_2 \hookrightarrow \mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1} \hookrightarrow \mathbb{H}_{-2}$ be a system of densely embedded Hilbert spaces. The norm in $\mathbb{H}_k$ is denoted by $|\cdot|_k$, the scalar product in $\mathbb{H}_0$ and the duality between $\mathbb{H}_1$ and $\mathbb{H}_{-1}$, as well as $\mathbb{H}_2$ and $\mathbb{H}_{-2}$ is denoted by $(\cdot\,,\cdot)$. Consider stochastic processes y, v_n with values in $\mathbb{H}_2$, $\mathbb{H}_{-1}$, respectively. Assume that processes y, v_n, have stochastic differentials

$$\begin{aligned} dy(t) &= a(t)dt + b^j(t)dW^j(t), \\ dv_n(t) &= f_n(t)dt + g_n^j(t)dW_n^j(t), \end{aligned}$$

where W is an r-dimensional Wiener process, and W_n its approximation satisfying Assumption 3.3.1. Assume moreover that g_n^j has a differential

$$dg_n^j(t) = p_n^j(t)dt + q_n^{jl}(t)dW_n^l(t).$$

Suppose that processes y, a, b^j, v_n f_n, g_n^j, p_n^j, q_n^{jl}, $j, l = 1, 2, \ldots, r$ satisfy the following conditions,

$$\begin{aligned} y, b^j &= O(1) \text{ in } \mathbb{H}_2, & v_n, g_n^j &= O(1) \text{ in } \mathbb{H}_{-1}, & q_n^{jl} &= O(1) \text{ in } \mathbb{H}_{-2}, \\ \int_0^T |a(s)|_2^2 ds &= O(1), & \int_0^T |f_n(s)|_{-1}^2 ds &= O(1), & \int_0^T |p_n^j(s)|_{-2}^2 ds &= O(1). \end{aligned}$$

Let α be a positive number, and $\{\psi_n\}_{n\in\mathbb{N}}$ a sequence of real valued stochastic processes, such that for every $\kappa < \alpha$

$$\psi_n = O(n^{-\kappa}),$$

and for every $\delta > 0$

$$||\psi_n||(T) = O(\ln^\delta n).$$

Lemma 3.4.4. *Under the assumption above for every $\beta < \alpha$*

$$\int_0^t (v_n(s), y(s))d\psi_n(s) = O(n^{-\beta})$$

providing that this integral exists.

Proof. Let us write the assumptions of the lemma in the form

$$\begin{aligned} &\sup_{t\le T} |y(t)|_2 \le \zeta, & &\sup_{t\le T} |v_n(t)|_{-1} \le \zeta, & & \\ &\int_0^T |a(t)|_2^2 dt \le \zeta, & &\int_0^T |f_n(t)|_{-1}^2 dt \le \zeta, & &\int_0^T |p_n^j(t)|_{-2}^2 dt \le \zeta, \\ &\sup_{t\le T} |b^j(t)|_2 \le \zeta, & &\sup_{t\le T} |g_n^j(t)|_{-1} \le \zeta, & &\sup_{t\le T} |q_n^{jl}(t)|_{-2} \le \zeta, \end{aligned}$$

for an a.s. finite random variable ζ,

$$||\psi_n||(T) \le \eta_\delta \ln^\delta n,$$

$$\sup_{t\le T} |\psi_n(t)| \le \xi_\kappa n^{-\kappa}$$

for all $\delta > 0$, positive $\kappa < \alpha$ and an a.s. for every δ, κ finite random variables η_δ, ξ_κ.

Let us choose some positive β, κ such that $\beta < \kappa < \alpha$, and fix any positive δ. We rewrite $(v_n(t), y(t))d\psi_n(t)$ as $(v_n(t), d(y \cdot \psi_n(t)))$. Here $\alpha \cdot \beta(t)$ denotes the integral $\int_0^t \alpha(s)d\beta(s)$. Integration by parts gives

$$(v_n, y) \cdot \psi_n(t) = \sum_{k=1}^{3} \theta_n^k(t) - \int_0^t (g_n^j(s), y \cdot \psi_n(s))d(W_n^j(s) - W^j(s)),$$

where

$$\begin{aligned}
\theta_n^1(t) &= (v_n(t), y \cdot \psi_n(t)) - (v_n(0), y \cdot \psi_n(0)), \\
\theta_n^2(t) &= -\int_0^t (f_n(s), y \cdot \psi_n(s))ds, \\
\theta_n^3(t) &= -\int_0^t (g_n^j(s), y \cdot \psi_n(s))dW^j(s).
\end{aligned}$$

Next,

$$-\int_0^t (g_n^j(s), y \cdot \psi_n(s))d(W_n^j(s) - W^j(s)) = \sum_{k=4}^{7} \theta_n^k(t),$$

where

$$\begin{aligned}
\theta_n^4(t) &= -(W_n^j(t) - W^j(t))(g_n^j(t), y \cdot \psi_n(t)) \\
&\quad + (W_n^j(0) - W^j(0))(g_n^j(0), y \cdot \psi_n(0)), \\
\theta_n^5(t) &= \int_0^t (W_n^j(s) - W^j(s))(p_n^j(s), y \cdot \psi_n(s))ds, \\
\theta_n^6(t) &= \int_0^t (q_n^{jl}(s), y \cdot \psi_n(s))dA_n^{jl}(s), \\
\theta_n^7(t) &= \int_0^t (W_n^j(s) - W^j(s))(g_n^j(s), y(s))d\psi_n(s),
\end{aligned}$$

and the process

$$A_n^{jl}(t) = \int_0^t (W_n^j(s) - W^j(s))dW_n^l(s)$$

by the assumptions imposed on approximation W_n is of bounded variation.

Let us show that for any $\kappa < \alpha$

$$\sup_t |y \cdot \psi_n(t)|_2 \leq \mu_\kappa n^{-\kappa}, \tag{3.4.4}$$

where μ_κ is a finite random variable. Indeed, assigning $\mathbb{V} = \mathbb{H}_2$ stochastic processes y, a, b^j, ψ_n satisfy the assumptions of Lemma 3.4.3, and therefore (3.4.4) holds.

Bellow we estimate terms $\theta_n^k(t)$.

$$|\theta_n^1(t)| \leq 2\sup_t |v_n(t)|_{-1} \sup_t |y \cdot \psi_n(t)|_1 \leq N\zeta\mu_\kappa n^{-\kappa},$$

$$|\theta_n^2(t)| \leq \left(\int_0^t |f_n(s)|_{-1}^2 ds\right)^{1/2} \left(\int_0^t |y \cdot \psi_n(s)|_1^2 ds\right)^{1/2} \leq N\sqrt{T}\zeta\mu_\kappa n^{-\kappa}.$$

Assign $\varphi_n(t) = (g_n^j(t), y \cdot \psi_n(t))$. By (3.4.4) and assumptions of the lemma $\sup_t |\varphi_n(t)| \leq \sup_t |g_n^j(t)|_{-1} |y \cdot \psi_n(t)|_1 \leq N\zeta\mu_\kappa n^{-\kappa}$ for $\kappa < \alpha$. Hence, by Lemma 3.4.2 for $\kappa < \alpha$

$$|\theta_n^3(t)| \leq \eta_\kappa n^{-\kappa},$$

where random variable η_κ is almost surely finite. Next,

$$|\theta_n^4(t)| \leq 2\sup_t |W_n^j(t) - W^j(t)| \sup_t |g_n^j(t)|_{-1} \sup_t |y \cdot \psi_n(t)|_1 \leq N\zeta\mu_\kappa n^{-2\kappa},$$

$$\begin{aligned} |\theta_n^5(t)| &\leq \sup_t |W_n^j(t) - W^j(t)| \left(\int_0^t |p_n^j(s)|_{-2}^2 ds\right)^{1/2} \left(\int_0^t |y \cdot \psi_n(s)|_2^2 ds\right)^{1/2} \\ &\leq \sqrt{T}\zeta\mu_\kappa n^{-2\kappa}, \end{aligned}$$

$$|\theta_n^6(t)| \leq ||A_n^{jl}||(T) \sup_t |q_n^{jl}(t)|_{-2} \sup_t |y \cdot \psi_n(t)|_2 \leq \vartheta_\kappa \zeta \eta_\kappa n^{-\kappa},$$

where random variable ϑ_κ is an almost surely finite by Assumption 3.3.1. Finally,

$$\begin{aligned} |\theta_n^7(t)| &\leq ||\psi_n||(T) \sup_t |W_n^j(t) - W^j(t)| \sup_t |g_n^j(t)|_{-1} \sup_t |y(t)|_1 \\ &\leq N\zeta^2 \eta_\delta (\ln n)^\delta n^{-\kappa}. \end{aligned}$$

It is easy to show that this implies $|\theta_n^7(t)| \leq Cn^{-\beta}$ for some constant C depending on β, κ and δ but independent of n. It suffices to sum up the expressions in the right-hand sides of the seven inequalities above, so that we get an almost surely finite random variable independent of n times $n^{-\beta}$. □

Lemma 3.4.5 (Gronwall). *Assume that for non-negative increasing continuous processes y, Q such that $Ey_\infty < \infty$ and $Q_\infty \leq K$ the following condition holds*

$$Ey_\tau \leq E\int_0^\tau y_s dQ_s + \varepsilon$$

for any stopping time τ. Then

$$Ey_\infty \leq \varepsilon e^K.$$

The next lemma is a modification of Gronwall lemma.

Lemma 3.4.6. *Consider a real bounded non-negative increasing continuous processes y_n. Assume that almost surely*

$$y_n(t) \leq \int_0^t y_n(s)dQ_n(s) + m_n(t) + q_n(t),$$

for any $t \in [0,T]$, where Q_n is a non-negative increasing continuous processes such that for any $\delta > 0$

$$Q_n = O(\ln^\delta n).$$

Suppose that

$$m_n(t) = \int_0^t v_n^j(s)dW^j(s),$$

where W is an r-dimensional Wiener process, and for some a.s. finite random variable ζ for every $t \in [0,T]$

$$|v_n^j(t)| \leq \zeta y_n(t).$$

Assume also that for every positive $\kappa < \alpha$

$$q_n = O(n^{-\kappa}).$$

Then

$$y_n = O(n^{-\gamma})$$

for all positive $\gamma < \alpha$.

Proof. Let us fix any arbitrary positive numbers γ, β, κ such that $\gamma < \beta < \kappa < \alpha$, and choose $r > (\beta - \gamma)^{-1}$ and $\delta < 1/r$. Denote

$$G_n(t) = \frac{Q_n(t)}{\ln^\delta n} + n^\kappa q_n(t) + \zeta^r t,$$

and for a positive number R define stopping time

$$\tau_R^n = \inf\{t \geq 0,\ G_n(t) \geq R\}.$$

Then for every $t \leq T$

$$\begin{aligned}|y_n(t \wedge \tau_R^n)|^r &\leq c_1(R\ln^\delta n)^{r-1}\int_0^t |y_n(s)|^r dQ_n(s \wedge \tau_R^n) \\ &\quad + c_1|m_n(t \wedge \tau_R^n)|^r + c_1 n^{-\kappa r},\end{aligned}$$

where constant $c_1 = c_1(r,R)$ does not depend on n. Taking supremum first of the r.h.s. and then of the l.h.s., and taking the expectation, we get for every t

$$\begin{aligned}E\sup_{s \leq t\wedge\tau_R^n\wedge T}|y_n(s)|^r &\leq c_1(R\ln^\delta n)^{r-1}E\int_0^t |y_n(s)|^r dQ_n(s \wedge \tau_R^n \wedge T) \\ &\quad + c_1 E\sup_{s \leq t\wedge\tau_R^n\wedge T}|m_n(s)|^r + c_1 n^{-\kappa r}.\end{aligned}$$

By Burkholder inequality and the definition of τ_R^n

$$E \sup_{s\le t\wedge\tau_R^n\wedge T}\left|\int_0^s v_n^j(u)dW^j(u)\right|^r \;\le\; c_2 E\int_0^t |y_n(s)|^r d((s\wedge\tau_R^n\wedge T)\zeta^r),$$

for some constant $c_2 = c_2(r)$. Therefore, denoting

$$x_n(t) \;=\; \sup_{s\le t\wedge\tau_R^n\wedge T} |y_n(s)|^r,$$

$$Q_n^{(1)}(t) \;=\; c_3((t\wedge\tau_R^n\wedge T)\zeta^r \;+\; (R\ln^\delta n)^{r-1}Q_n(t\wedge\tau_R^n\wedge T))$$

for some appropriate constant $c_3 = c_3(r,R)$, we get

$$Ex_n(\tau) \;\le\; E\int_0^\tau x_n(s)dQ_n^{(1)}(s) \;+\; c_1 n^{-\kappa r}$$

for any stopping time τ. Since $Q_n^{(1)}(t) \le c_3(R+(R\ln^\delta n)^r)$, by Gronwall Lemma 3.4.5

$$E\sup_{s\le\tau_R^n\wedge T}|y_n(s)|^r \;\le\; c_1 n^{-\kappa r}\exp(c_3(R+(R\ln^\delta n)^r)).$$

This, under condition $\delta r < 1$, implies for every positive $\beta < \kappa$, in particular for β chosen above

$$E\sup_{s\le\tau_R^n\wedge T}|y_n(s)|^r \;\le\; c_\beta n^{-\beta r},$$

where constant $c_\beta = c_\beta(r,R)$ does not depend on n. Then by Lemma 3.4.1

$$\sup_{s\le\tau_R^n\wedge T} y_n(s) \;\le\; \eta_{\gamma,R}n^{-\gamma},$$

where random variable $\eta_{\gamma,R}$ is a.s. finite.

Define stochastic process $G(t) = \sup_n G_n(t)$. Then stopping time

$$\tau_R \;=\; \inf\{t\ge 0,\; G(t)\ge R\}$$

does not exceed τ_R^n, and by the last inequality

$$\sup_{s\le\tau_R\wedge T} y_n(s) \;\le\; \eta_{\gamma,R}n^{-\gamma}.$$

By assumptions of the lemma process $G(t)$ is bounded on the interval $[0,T]$ by an a.s. finite random variable. This implies a.s.

$$\lim_{R\to\infty}\tau_R \;=\; \infty.$$

Consider sets $\Omega_R = \{\omega,\; \tau_R\ge T\}$. Obviously, a.e. ω hits some Ω_R, so that $\mathbf{P}(\cup_{R=1}^\infty\Omega_R) = 1$. Define $\eta_\gamma = \eta_{\gamma,R}$ for $\omega\in\Omega_R\backslash(\cup_{k=1}^{R-1}\Omega_k)$, $R = 1,2,\ldots$, and $\eta_\gamma = \infty$ for those values of ω that do not hit any Ω_R. Then η_γ is an a.s. finite random variable, and

$$\sup_{s\le T} y_n(s) \;\le\; \eta_\gamma n^{-\gamma}.$$

□

3.5 Proof of Theorem 3.3.1

The central tool of the proof is the Gronwall-type Lemma 3.4.6. We transform equations (3.3.1), (3.3.2) to an inequality of a special form so that we can make use of the lemma.

In order to simplify notations let us introduce for every $j = 1, 2, \ldots, r$, $l = 0, 1, \ldots, r$ operators, non-linear in general,

$$\begin{aligned} \mathbb{A}(v) &= Av + f, \\ \mathbb{B}^j(v) &= B^j v + g^j, \\ \mathbb{B}^{j(l)}(v) &= B^{j(l)} v + g^{j(l)}, \end{aligned}$$

and for every $n = 1, 2, \ldots$

$$\begin{aligned} \mathbb{A}_n(v) &= A_n v + f_n, \\ \mathbb{B}^j_n(v) &= B^j_n v + g^j_n, \\ \mathbb{B}^{j(l)}_n(v) &= B^{j(l)}_n v + g^{j(l)}_n. \end{aligned}$$

Note that by Assumption 3.3.5 for every $\kappa < \alpha$ and $j = 1, 2, \ldots, r$, $l = 0, 1, \ldots, r$

$$\begin{aligned} \mathbb{A} - \mathbb{A}_n &= O(n^{-\kappa}) \ in\ \mathbb{L}(\mathbb{H}_2, \mathbb{H}_0), \\ \mathbb{B}^j - \mathbb{B}^j_n &= O(n^{-\kappa}) \ in\ \mathbb{L}(\mathbb{H}_2, \mathbb{H}_1), \\ \mathbb{B}^{j(l)} - \mathbb{B}^{j(l)}_n &= O(n^{-\kappa}) \ in\ \mathbb{L}(\mathbb{H}_1, \mathbb{H}_0). \end{aligned} \tag{3.5.1}$$

Equations (3.3.1), (3.3.2) can be rewritten as

$$\begin{aligned} du &= \mathbb{A}(u)dt + \mathbb{B}^j(u)dW^j + \frac{1}{2}B^j\mathbb{B}^j(u)dt + \frac{1}{2}\mathbb{B}^{l(l)}(u)dt, \\ u(0) &= \xi, \end{aligned} \tag{3.5.2}$$

$$\begin{aligned} du_n &= \mathbb{A}_n(u_n)dt + \mathbb{B}^j_n(u_n)dW^j_n, \\ u_n(0) &= \xi_n. \end{aligned} \tag{3.5.3}$$

Using the Itô formula for $|u - u_n|$ in the triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$ we get for every $t \in [0, T]$

$$\begin{aligned} d|u - u_n|^2 &= 2(u - u_n, A_n(u - u_n))dt \\ &\quad + 2(u - u_n, (\mathbb{A} - \mathbb{A}_n)(u))dt \\ &\quad + 2(u - u_n, B^j_n(u - u_n))dW^j \\ &\quad + 2(u - u_n, \mathbb{B}^j_n(u_n))d(W^j - W^j_n) \\ &\quad + 2(u - u_n, (\mathbb{B}^j - \mathbb{B}^j_n)(u))dW^j \\ &\quad + (u - u_n, B^j\mathbb{B}^j(u))dt \\ &\quad + (u - u_n, \mathbb{B}^{l(l)}(u))dt \\ &\quad + |\mathbb{B}^j(u)|^2 dt. \end{aligned}$$

Next, we rewrite the third term in the previous expansion. We take separately differentials of $B^j_n u_n$ and g^j_n. For every $v \in \mathbb{H}_1$

$$\begin{aligned} d(v, B^j_n u_n) &= (v, B^{j(0)}_n u_n)dt \\ &+ (v, B^{j(l)}_n u_n)dW^l_n \\ &+ (v, B^j_n \mathbb{A}_n(u_n))dt \\ &+ (v, B^j_n \mathbb{B}^l_n(u_n))dW^l_n, \end{aligned}$$

$$\begin{aligned} d(v, g^j_n) &= (v, g^{j(0)}_n)dt \\ &+ (v, g^{j(l)}_n)dW^l_n. \end{aligned}$$

Summing up the results we get,

$$\begin{aligned} d(v, \mathbb{B}^j_n(u_n)) &= (v, \mathbb{B}^{j(0)}_n(u_n))dt \\ &+ (v, \mathbb{B}^{j(l)}_n(u_n))dW^l_n \\ &+ (v, B^j_n \mathbb{A}_n(u_n))dt \\ &+ (v, B^j_n \mathbb{B}^l_n(u_n))dW^l_n. \end{aligned}$$

Integration by parts, the Itô formula in the triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \hookrightarrow \mathbb{H}_{-1}$ and the previous identity give

$$\begin{aligned} &(u - u_n, \mathbb{B}^j_n(u_n))d(W^j - W^j_n) \\ &= d\{(W^j - W^j_n)(u - u_n, \mathbb{B}^j_n(u_n))\} \\ &- (W^j - W^j_n)(A_n(u - u_n), \mathbb{B}^j_n(u_n))dt \\ &- (W^j - W^j_n)((\mathbb{A} - \mathbb{A}_n)(u), \mathbb{B}^j_n(u_n))dt \\ &- (W^j - W^j_n)(\mathbb{B}^l(u), \mathbb{B}^j_n(u_n))dW^l \\ &+ (W^j - W^j_n)(\mathbb{B}^l_n(u_n), \mathbb{B}^j_n(u_n))dW^l_n \\ &- \frac{1}{2}(W^j - W^j_n)(B^l \mathbb{B}^l(u), \mathbb{B}^j_n(u_n))dt \\ &- \frac{1}{2}(W^j - W^j_n)(\mathbb{B}^{l(l)}(u), \mathbb{B}^j_n(u_n))dt \\ &- (W^j - W^j_n)(u - u_n, \mathbb{B}^{j(0)}_n(u_n))dt \\ &- (W^j - W^j_n)(u - u_n, \mathbb{B}^{j(l)}_n(u_n))dW^l_n \\ &- (W^j - W^j_n)(u - u_n, B^j_n \mathbb{A}_n(u_n))dt \\ &- (W^j - W^j_n)(u - u_n, B^j_n \mathbb{B}^l_n(u_n))dW^l_n \\ &- (\mathbb{B}^j(u), \mathbb{B}^j_n(u_n))dt. \end{aligned}$$

It is easy to verify that the following three equalities hold for every $t \in [0, T]$,

$$
\begin{aligned}
&|\mathbb{B}^j(u)|^2 dt \\
&- 2(\mathbb{B}^j(u), \mathbb{B}^j_n(u_n))dt \\
&+ 2(W^j - W^j_n)(\mathbb{B}^l_n(u_n), \mathbb{B}^j_n(u_n))dW^l_n \\
&\qquad = 2(B^j_n(u - u_n), B^l_n(u - u_n))(W^j - W^j_n)dW^l_n \\
&\qquad\quad + 2d(J^{(1)}_n + K^{(1)}_n + K^{(2)}_n + \rho^{(1)}_n),
\end{aligned}
$$

$$
\begin{aligned}
&(u - u_n, B^j \mathbb{B}^j(u))dt \\
&- 2(W^j - W^j_n)(u - u_n, B^j_n \mathbb{B}^l_n(u_n))dW^l_n \\
&\qquad = 2(u - u_n, B^j_n B^l_n(u - u_n))(W^j - W^j_n)dW^l_n \\
&\qquad\quad + 2d(J^{(2)}_n + K^{(3)}_n + \rho^{(2)}_n),
\end{aligned}
$$

$$
\begin{aligned}
&(u - u_n, \mathbb{B}^{l(l)}(u))dt \\
&- 2(W^j - W^j_n)(u - u_n, \mathbb{B}^{j(l)}_n(u_n))dW^l_n \\
&\qquad = 2(u - u_n, B^{j(l)}_n(u - u_n))(W^j - W^j_n)dW^l_n \\
&\qquad\quad + 2d(J^{(3)}_n + K^{(4)}_n + \rho^{(3)}_n),
\end{aligned}
$$

where

$$
\begin{aligned}
J^{(1)}_n &= -\int_0^t (\mathbb{B}^j(u), \mathbb{B}^l(u))dS^{jl}_n, \\
J^{(2)}_n &= -\int_0^t (u, B^j \mathbb{B}^l(u))dS^{jl}_n, \\
J^{(3)}_n &= -\int_0^t (u, \mathbb{B}^{j(l)}(u))dS^{jl}_n,
\end{aligned}
$$

$$
\begin{aligned}
K^{(1)}_n &= \int_0^t (\mathbb{B}^j(u_n), \mathbb{B}^l(u))dS^{jl}_n, \\
K^{(2)}_n &= \int_0^t (\mathbb{B}^j(u), \mathbb{B}^l(u_n))dS^{jl}_n, \\
K^{(3)}_n &= \int_0^t (u_n, B^j \mathbb{B}^l(u))dS^{jl}_n, \\
K^{(4)}_n &= \int_0^t (u_n, \mathbb{B}^{j(l)}(u))dS^{jl}_n,
\end{aligned}
$$

$$
\begin{aligned}
\rho_n^{(1)}(t) &= \int_0^t \{((\mathbb{B}^j(u), \mathbb{B}^l(u)) - (\mathbb{B}_n^j(u), \mathbb{B}_n^l(u))) \\
&\qquad + (\mathbb{B}_n^j(u_n), (\mathbb{B}_n^l - \mathbb{B}^l)(u)) \\
&\qquad + (\mathbb{B}_n^l(u_n), (\mathbb{B}_n^j - \mathbb{B}^j)(u))\} dA_n^{jl} \\
&\quad + \int_0^t \{(\mathbb{B}^j(u), (\mathbb{B}_n^l - \mathbb{B}^l)(u_n)) \\
&\qquad + (\mathbb{B}^l(u), (\mathbb{B}_n^j - \mathbb{B}^j)(u_n))\} dS_n^{jl}, \\
\rho_n^{(2)}(t) &= \int_0^t (u - u_n, (B^j\mathbb{B}^l - B_n^j\mathbb{B}_n^l)(u)) dA_n^{jl}, \\
\rho_n^{(3)}(t) &= \int_0^t (u - u_n, (\mathbb{B}^{j(l)} - \mathbb{B}_n^{j(l)})(u)) dA_n^{jl}.
\end{aligned}
$$

Here we denoted

$$
A_n^{jl}(t) = \int_0^t (W^j(s) - W_n^j(s)) dW_n^l(s).
$$

Note that by Assumption 3.3.1 for each $\delta > 0$

$$
||A_n^{jl}|| = O(\ln^\delta n).
$$

We derive from the identities above

$$
|u - u_n|^2 = |\xi - \xi_n|^2 + 2\sum_{k=1}^{5} I_n^{(k)} + 2q_n, \tag{3.5.4}
$$

where

$$
\begin{aligned}
I_n^{(1)} &= \int_0^t (u - u_n, A_n(u - u_n)) ds, \\
I_n^{(2)} &= \int_0^t (u - u_n, B_n^j(u - u_n)) dW^j, \\
I_n^{(3)} &= \int_0^t (B_n^j(u - u_n), B_n^l(u - u_n)) dA_n^{jl}, \\
I_n^{(4)} &= \int_0^t (u - u_n, B_n^j B_n^l(u - u_n)) dA_n^{jl}, \\
I_n^{(5)} &= \int_0^t (u - u_n, B_n^{j(l)}(u - u_n)) dA_n^{jl},
\end{aligned}
$$

$$
q_n = p_n + r_n + k_n + l_n + m_n + \sum_{k=1}^{3} \rho_n^{(k)} + \sum_{k=1}^{3} J_n^{(k)} + \sum_{k=1}^{4} K_n^{(k)},
$$

$$
\begin{aligned}
p_n &= (u(t) - u_n(t), \mathbb{B}_n^j(u_n)(t))(W^j(t) - W_n^j(t)) \\
&\quad - (u(0) - u_n(0), \mathbb{B}_n^j(u_n)(0))(W^j(0) - W_n^j(0)),
\end{aligned}
$$

$$
\begin{aligned}
r_n &= \int_0^t (u - u_n, (\mathbb{A} - \mathbb{A}_n)(u))ds \\
&+ \int_0^t (u - u_n, (\mathbb{B}^j - \mathbb{B}^j_n)(u))dW^j,
\end{aligned}
$$

$$
\begin{aligned}
k_n &= \int_0^t (W^j - W^j_n)(A_n u_n, B^j_n u_n)ds \\
&+ \int_0^t (W^j - W^j_n)(u_n, B^j_n A_n u_n)ds,
\end{aligned}
$$

$$
\begin{aligned}
l_n = &- \int_0^t (W^j - W^j_n)(A_n u, \mathbb{B}^j_n(u_n))ds \\
&- \int_0^t (W^j - W^j_n)((\mathbb{A} - \mathbb{A}_n)(u), \mathbb{B}^j_n(u_n))ds \\
&- \frac{1}{2}\int_0^t (W^j - W^j_n)(B^l \mathbb{B}^l(u), \mathbb{B}^j_n(u_n))ds \\
&- \frac{1}{2}\int_0^t (W^j - W^j_n)(\mathbb{B}^{l(l)}(u), \mathbb{B}^j_n(u_n))ds \\
&- \int_0^t (W^j - W^j_n)(u - u_n, \mathbb{B}^{j(0)}_n(u_n))ds \\
&- \int_0^t (W^j - W^j_n)(u, B^j_n \mathbb{A}_n(u_n))ds \\
&+ \int_0^t (W^j - W^j_n)(A_n u_n, g^j_n)ds \\
&+ \int_0^t (W^j - W^j_n)(u_n, B^j_n f_n)ds,
\end{aligned}
$$

$$
m_n = \int_0^t (W^j - W^j_n)(\mathbb{B}^l(u), \mathbb{B}^j_n(u_n))dW^l.
$$

Let us introduce process

$$
y_n(t) = |u(t) - u_n(t)|_0^2 + \lambda \int_0^t |u(s) - u_n(s)|_1^2 ds,
$$

and define stopping time

$$
\pi_{n,\varepsilon} = \inf\{t \geq 0,\ y_n(t) \geq \varepsilon\}.
$$

By Assumption 3.3.6 process $|u|_0$ is bounded by an a.s. finite random variable uniformly in t on the interval $[0, T]$. Then processes $|u_n|_0$ is bounded uniformly in n uniformly in t on the interval $[0, \pi_{n,\varepsilon} \wedge T]$. By the same argument the boundedness by an a.s. finite random variable for $\int_0^T |u|_1 ds$ implies the boundedness by an a.s. finite random variable uniformly in n for $\int_0^{\pi_{n,\varepsilon}\wedge T}|u_n|_1 ds$. So that for some finite random variable μ we have,

$$
\begin{aligned}
|u(t)|_k &\le \mu, & |u_n(t\wedge\pi_{n,\varepsilon})|_0 &\le \mu, \\
\int_0^T |u(s)|_l^2 ds &\le \mu, & \int_0^{\pi_{n,\varepsilon}\wedge T} |u_n(s)|_1^2 ds &\le \mu.
\end{aligned} \tag{3.5.5}
$$

for every $t \in [0,T]$ and any integers $k \le 5$, $l \le 6$.

Next, we rewrite equality (3.5.4). By Assumption 3.3.4

$$
I_n^{(1)}(t) \le \zeta \int_0^t |u(s)-u_n(s)|_0^2 ds - \lambda \int_0^t |u(s)-u_n(s)|_1^2 ds,
$$

$$
|I_n^{(3)}(t)+I_n^{(4)}(t)| + |I_n^{(5)}(t)| \le 2\zeta \int_0^t |u(s)-u_n(s)|_0^2 d||A_n^{jl}||(s).
$$

For every ε let us introduce

$$
y_{n\varepsilon}(t) = y_n(t\wedge\pi_{n,\varepsilon}),
$$

and similarly define $q_{n\varepsilon}$, $J_{n\varepsilon}^{(k)}$, $p_{n\varepsilon}$, $r_{n\varepsilon}$, $k_{n\varepsilon}$, $l_{n\varepsilon}$, $m_{n\varepsilon}$, $\rho_{n\varepsilon}^{(k)}$. Taking into account inequalities above it follows from (3.5.4) that for every $\varepsilon > 0$ and $t \in [0,T]$

$$
y_{n\varepsilon}(t) \le \int_0^t y_{n\varepsilon}(s) dQ_{n\varepsilon}(s) + 2I_n^{(2)}(t\wedge\pi_{n,\varepsilon}) + 2q_{n\varepsilon}(t),
$$

where

$$
Q_{n\varepsilon}(t) = 2\zeta(t\wedge\pi_{n,\varepsilon} + ||A_n^{jl}||(t\wedge\pi_{n,\varepsilon}))
$$

is a non-negative increasing continuous processes, and $Q_{n\varepsilon} = O(\ln^\delta n)$ for every $\delta > 0$. Our aim now is to show that for every $\kappa < \alpha$ and $\varepsilon > 0$

$$
q_{n\varepsilon} = O(n^{-\kappa}). \tag{3.5.6}
$$

Then we are in position to apply Lemma 3.4.6. Let us show that $p_{n\varepsilon}$ satisfies this property (i.e. (3.5.6) holds if we replace $q_{n\varepsilon}$ by $p_{n\varepsilon}$). Indeed,

$$
|(u-u_n, \mathbb{B}_n^j(u_n))| \le |(u-u_n, B_n^j(u-u_n))| + |(u-u_n, \mathbb{B}_n^j(u))|,
$$

and by Assumptions 3.3.3, 3.3.4

$$
(u-u_n, \mathbb{B}_n^j(u_n)) = O(1).
$$

Then Assumption 3.3.1 implies for every $\kappa < \alpha$

$$
\sup_{t\le T} |p_{n\varepsilon}(t)| \le \xi_\kappa n^{-\kappa}
$$

for some a.s. finite random variable ξ_κ.

We move on to $k_{n\varepsilon}$. We have

$$|k_{n\varepsilon}(t)| \leq \sup_{t\leq T}|W^j - W_n^j| \int_0^T |(A_n u_n, B_n^j u_n) + (u_n, B_n^j A_n u_n)|_0 ds,$$

and by Assumptions 3.3.1, 3.3.4 for every $\varepsilon > 0$, $\kappa < \alpha$

$$|k_{n\varepsilon}(t)| \leq \eta_\kappa n^{-\kappa} \int_0^{T\wedge\pi_{n,\varepsilon}} |u_n|_1^2 ds,$$

for some a.s. finite random variable η_κ. This proves property (3.5.6) for $k_{n\varepsilon}$.

Next, we check (3.5.6) for $l_{n\varepsilon}$. By Assumption 3.3.1 for every $\varepsilon > 0$, $\kappa < \alpha$ and some a.s. finite random variable ν_κ

$$\begin{aligned}|l_{n\varepsilon}(t)| \leq \nu_\kappa n^{-\kappa} \int_0^{\pi_{n,\varepsilon}\wedge T} & \{|(A_n u, \mathbb{B}_n^j(u_n))| + |((\mathbb{A} - \mathbb{A}_n)(u), \mathbb{B}_n^j(u_n))| \\ & + |(B^l \mathbb{B}^l(u), \mathbb{B}_n^j(u_n))| + |(\mathbb{B}^{l(l)}(u), \mathbb{B}_n^j(u_n))| \\ & + |(u - u_n, \mathbb{B}_n^{j(0)}(u_n))| + |(u, B_n^j \mathbb{A}_n(u_n))| \\ & + |(A_n u_n, g_n^j)| + |(u_n, B_n^j f_n)|\} ds.\end{aligned}$$

It is easy to show that the integral in the r.h.s. is uniformly in n bounded by an a.s. finite random variable. For example, by Assumption 3.3.3 and (3.5.5) it is true for the first term,

$$\begin{aligned}&\int_0^{\pi_{n,\varepsilon}\wedge T} |(A_n u, \mathbb{B}_n^j(u_n))| ds \\ &\leq \left\{|A_n|_{2,0}^2 \int_0^T |u|_2^2 ds\right\}^{1/2} \left\{|B_n^j|_{1,0}^2 \int_0^{\pi_{n,\varepsilon}\wedge T} |u_n|_1^2 ds + \int_0^T |g_n^j|_0^2 ds\right\}^{1/2}.\end{aligned}$$

All other terms are estimated in the same way.

To show that $m_{n\varepsilon}$ satisfies (3.5.6) we assign $\varphi_n(t) = (W^j - W_n^j)(\mathbb{B}^l(u), \mathbb{B}_n^j(u_n))$. By Assumption 3.3.3 and (3.5.5)

$$|(\mathbb{B}^l(u), \mathbb{B}_n^j(u_n))| \leq (|B^l|_{2,1}|u|_2 + |g^l|_1)(|B_n^j|_{0,-1}|u_n|_0 + |g_n^j|_{-1}),$$

which can be estimated by an a.s. finite random variable. Hence, by Assumption 3.3.1 $\varphi_n = O(n^{-\kappa})$ for $\kappa < \alpha$, and then by Lemma 3.4.2 so does $m_{n\varepsilon}$ for every $\varepsilon > 0$.

Next, we consider $r_{n\varepsilon}$. We have

$$\begin{aligned}&\int_0^{\pi_{n,\varepsilon}\wedge T} |(u - u_n, (\mathbb{A} - \mathbb{A}_n)(u))| ds \\ &\leq \sup_{t\leq T} |(A - A_n)u|_{-1} \left\{\int_0^{\pi_{n,\varepsilon}\wedge T} |u - u_n|_1^2 ds\right\}^{1/2} \\ &\quad + T \sup_{t\leq \pi_{n,\varepsilon}\wedge T} |u - u_n|_0 |f - f_n|_0\end{aligned}$$

which by Assumption 3.3.5 and the definition of $\pi_{n,\varepsilon}$ can be estimated by an a.s. finite random variable. To estimate the second term in $r_{n\varepsilon}$ we produce

$$|(u-u_n,(\mathbb{B}^j-\mathbb{B}^j_n)(u))| \;\le\; |u-u_n|_0(|(B^j-B^j_n)u|_0+|g^j-g^j_n|_0),$$

i.e. $(u-u_n,(\mathbb{B}^j-\mathbb{B}^j_n)(u))=O(n^{-\kappa})$ for all $\kappa<\alpha$ and then apply Lemma 3.4.2.

Using the same arguments it is easy to show that $\rho^k_{n\varepsilon}$, $k=1,2,3$ satisfy property (3.5.6).

To prove that $J^{(k)}_{n\varepsilon}$, $k=1,2,3$, satisfy (3.5.6) we use Lemma 3.4.3. Let us show this for $J^{(1)}_{n\varepsilon}$. Assign

$$\begin{aligned} y(t) &= (\mathbb{B}^j(u(t)),\mathbb{B}^l(u(t))),\\ \psi_n(t) &= S^{jl}_n(t). \end{aligned}$$

Then taking differential of $y(t)$, by the Itô formula for $(\mathbb{B}^j(u),\mathbb{B}^l(u))=(B^ju,B^lu)+2(B^ju,g^l)+(g^j,g^l)$ in the triple $\mathbb{H}_1\hookrightarrow\mathbb{H}_0\hookrightarrow\mathbb{H}_{-1}$ we get,

$$\begin{aligned} a(t) &= 2(B^j\mathbb{A}(u),\mathbb{B}^l(u)) \;+\; (B^jB^h\mathbb{B}^h(u),\mathbb{B}^l(u))\\ &\quad+\;(B^j\mathbb{B}^{h(h)}(u),\mathbb{B}^l(u)) \;+\; (\mathbb{B}^{j(h)}(u),\mathbb{B}^{l(h)}(u))\\ &\quad+\;(B^j\mathbb{B}^h(u),B^l\mathbb{B}^h(u)) \;+\; 2(\mathbb{B}^{j(h)}(u),B^l\mathbb{B}^h(u)),\\ b^h(t) &= 2(\mathbb{B}^{j(h)}(u),\mathbb{B}^l(u)) \;+\; 2(B^j\mathbb{B}^h(u),\mathbb{B}^l(u)). \end{aligned}$$

By Assumptions 3.3.1, 3.3.3 the conditions of Lemma 3.4.3 with $\mathbb{V}=\mathbb{R}$ are satisfied. This lemma completes the proof for $J^{(1)}_{n\varepsilon}$. The proof for $J^{(4)}_{n\varepsilon}$ and $J^{(6)}_{n\varepsilon}$ is performed in the same way.

We move on to $K^{(k)}_{n\varepsilon}$, $k=1,2,3,4$. For $K^{(1)}_{n\varepsilon}$ we rewrite $K^{(1)}_n=\tilde{K}^{(1)}_n+J^{(K1)}_n$, where

$$\begin{aligned} \tilde{K}^{(1)}_n &= \int_0^t (u_n,B^{j*}\mathbb{B}^l(u))dS^{jl}_n,\\ J^{(K1)}_n &= \int_0^t (g^j,\mathbb{B}^l(u))dS^{jl}_n. \end{aligned}$$

To estimate $\tilde{K}^{(1)}_n$ we use Lemma 3.4.4. Assign

$$\begin{aligned} y(t) &= B^{j*}\mathbb{B}^l(u(t)),\\ v_n(t) &= u_n(t),\\ \psi_n(t) &= S^{jl}_n(t). \end{aligned}$$

Taking differentials of $y(t)$ and $v_n(t)$ we get

$$\begin{aligned} a(t) &= \frac{1}{2}B^{j*}B^l(2\mathbb{A}+B^h\mathbb{B}^h+\mathbb{B}^{h(h)})(u) \;+\; B^{j(h)*}(\mathbb{B}^{l(h)}+B^l\mathbb{B}^h)(u),\\ b^h(t) &= B^{j*}(\mathbb{B}^{l(h)}+B^l\mathbb{B}^h)(u) \;+\; B^{j(h)*}\mathbb{B}^l(u), \end{aligned}$$

$$
\begin{aligned}
f_n(t) &= \mathbb{A}_n(u_n), \\
g_n^h(t) &= \mathbb{B}_n^h(u_n),
\end{aligned}
$$

$$
\begin{aligned}
p_n^h(t) &= B_n^h \mathbb{A}_n(u_n), \\
q_n^{hk}(t) &= (\mathbb{B}_n^{h(k)} + B_n^h \mathbb{B}_n^k)(u_n).
\end{aligned}
$$

It is easy to check that by Assumptions 3.3.4, 3.3.6 all the functions above satisfy conditions of Lemma 3.4.4. To estimate $J_n^{(K1)}$ we assign

$$
\begin{aligned}
y(t) &= (g^j, \mathbb{B}^l(u)), \\
a(t) &= (g^j, B^l \mathbb{A}(u)) + \frac{1}{2}(g^j, B^l B^h \mathbb{B}^h(u)) + \frac{1}{2}(g^j, B^l \mathbb{B}^{h(h)}(u)) \\
&\quad + (g^{j(h)}, \mathbb{B}^{l(h)}(u)) + (g^{j(h)}, B^l \mathbb{B}^h(u)), \\
b^h(t) &= (g^{j(h)}, \mathbb{B}^l(u)) + (g^j, \mathbb{B}^{l(h)}(u)) + (g^j, B^l \mathbb{B}^h(u)).
\end{aligned}
$$

This functions and $\psi_n(t) = S_n^{jl}(t)$ satisfy the conditions of Lemma 3.4.3 with $\mathbb{V} = \mathbb{R}$. Thus, (3.5.6) holds for $K_n^{(1)}$ and, consequently, for $K_{n\varepsilon}^{(1)}$. The proof for $K_{n\varepsilon}^{(k)}$, $k = 2, 3, 4$ can be performed similarly.

Finally, as was mentioned above, we apply Lemma 3.4.6, and derive

$$
y_{n\varepsilon} = O(n^{-\gamma})
$$

for any $\varepsilon > 0$, $\gamma < \alpha$. To finish the proof of the theorem it suffices to apply Lemma 2.4.3.

□

Chapter 4

Stochastic Partial Differential Equations

4.1 Introduction

This chapter may be considered as a continuation of the previous one. Here we apply its main result, Theorem 3.3.1, to stochastic partial differential equation of the form

$$\begin{aligned} du(t,\omega,x) &= (\mathcal{L}u(t,\omega,x) + f(t,\omega,x))dt \\ &+ (\mathcal{M}_j u(t,\omega,x) + g_j(t,\omega,x))dW^j(t,\omega) \end{aligned} \tag{4.1.1}$$

in the normal triple of spaces $W_2^1 \hookrightarrow L_2 \equiv L_2^* \hookrightarrow {W_2^1}^*$, with Cauchy condition

$$u(0,\omega,x) = u_0(\omega,x), \tag{4.1.2}$$

where $\mathcal{L}$ is a second order elliptic differential operator, $\mathcal{M}_j$ is a first order differential operator for every $j = 1,2,\dots,d$; $f(t,\omega,x)$, $g_j(t,\omega,x)$ are random fields on $[0,T]\times\mathbb{R}^d$ for some positive T and $j = 1,2,\dots,d$, process W is a multidimensional Wiener process, and $u_0(\omega,x)$ is a random field on $\mathbb{R}^d$. Above and throughout the paper we use the summation convention with respect to the repeated indices.

Under the solution of the problem (4.1.1)-(4.1.2) we understand a continuous process u taking values in the Sobolev space W_2^m for some non-negative integer m and satisfying (4.1.1)-(4.1.2) in the generalized sense (see Definition 4.2.1 below).

We approximate the Wiener process W with a sequence of processes $\{W_n\}_{n\in\mathbb{N}}$ of bounded variation. The convergence is considered in the topology O with some rate of convergence (see page 29). Moreover, we approximate coefficients of the differential operators $\mathcal{L}$, $\mathcal{M}_j$ as well as the random fields f, g_j, u_0. We use the same topology with the same rate of convergence. Therefore, for every $n \in \mathbb{N}$ we have a second order partial differential equation

$$\begin{aligned} du_n(t,\omega,x) &= (\mathcal{L}_n u_n(t,\omega,x) + f_n(t,\omega,x))dt \\ &\quad + (\mathcal{M}_{nj} u_n(t,\omega,x) + g_{nj}(t,\omega,x))dW_n^j(t,\omega) \end{aligned} \tag{4.1.3}$$

with Cauchy condition

$$u_n(0,\omega,x) = u_{n0}(\omega,x). \tag{4.1.4}$$

Relying on the results of Theorem 3.3.1 we prove that under some natural conditions the solution u_n of the problem (4.1.3)-(4.1.4) converges in the topology O with the same rate of convergence. However, the limit $\tilde{u}$ is the solution of a stochastic partial differential equation related to the equation (4.1.1). It has an additional second order differential drift term. This equation can be considered as the equation (4.1.1) with the last differential written in the Stratonovich form.

The equations of the considered type are especially useful because of their applications to many important problems, in particular the problem of filtering for diffusion processes which can be reduced to the investigation of the equation of type (4.1.1). The filtering problem is studied in the next chapter.

4.2 Generalities

Before formulating the result, in this section we recall some fundamental concepts.

4.2.1 Sobolev Spaces

All considerations are carried out in Sobolev spaces. We recall some general ideas from the theory of these spaces. Let $\mathbb{R}^d$ be a d-dimensional Euclidean space. We fix an orthonormal basis in $\mathbb{R}^d$, and for $x \in \mathbb{R}^d$ we denote its coordinates by $x_1, \dots, x_d$, and its norm by $|x|$. For $p = 1, 2, \dots, d$ we denote by D_p the differential operator $\partial/\partial x_p$, and for $p = 0$ we assume D_0 to be the identity. A d-dimensional vector with non-negative integer components we call a *multi-index*. For a multi-index $\gamma = (\gamma_1, \dots, \gamma_d)$ of length $|\gamma| := \gamma_1 + \dots + \gamma_d$ we define D^γ as the differential operator

$$D^\gamma = D_1^{\gamma_1} D_2^{\gamma_2} \dots D_d^{\gamma_d}.$$

First, we recall the definitions of some classic spaces of functions. Suppose further that m is an integer, $m \geq 0$, $p \in (1, \infty)$. We denote by $C_0^\infty = C_0^\infty(\mathbb{R}^d)$ the space of real-valued infinitely differentiable functions on $\mathbb{R}^d$ with a finite support. By $C^n = C^n(\mathbb{R}^d)$ we denote the space of n times continuously differentiable functions on $\mathbb{R}^d$ with the finite norm

$$|f|_{C^n} = \sum_{|\gamma| \leq n} \sup_x |D^\gamma f(x)|,$$

and by $L_p = L_p(\mathbb{R}^d)$ we denote the space (of classes) of real-valued functions on $\mathbb{R}^d$ with the finite norm

$$|f|_{L_p} = \left(\int_{\mathbb{R}^d} |f(x)|^p dx\right)^{1/p}.$$

The space (of classes) of real-valued functions on $\mathbb{R}^d$ belonging together with their derivatives up to the order m to the space L_p is called *the Sobolev space* $W_p^m = W_p^m(\mathbb{R}^d)$. Sobolev space W_p^m equipped with the norm defined by

$$|f|_{W_p^m} = \left(\sum_{|\gamma|\le m} \int_{\mathbb{R}^d} |D^\gamma f(x)|^p dx\right)^{1/p}$$

is a separable Banach space. Moreover, for $p = 2$ the space W_2^m is a Hilbert space with respect to the scalar product $(\cdot\,,\cdot)_{W_2^m}$ generated by the norm $|\cdot|_{W_2^m}$.

For $p = 2$ there is another definition of Sobolev spaces. Let Δ be the Laplace operator $\sum_i \frac{\partial^2}{\partial x_i^2}$, I be the identity operator, and let us define operator $\Lambda = (I - \Delta)^{1/2}$. For $H_0 = L_2$ and operator Λ we can define Hilbert scale $\{H_\alpha\}_{\alpha\in\mathbb{R}}$ (see page 26): for $\alpha > 0$ we define H_α as the domain of operator Λ^α; for $\alpha \le 0$ H_α is the completion of space H_0 in the norm $|\cdot|_{H_\alpha} = |\Lambda^\alpha \cdot|_{L_2}$. Then for every $\alpha \in \mathbb{R}$ the space H_α is a Hilbert space with respect to the scalar product $(\cdot\,;\cdot)_{H_\alpha} = (\Lambda^\alpha\cdot\,,\Lambda^\alpha\cdot)_{L_2}$.

Proposition 4.2.1. *For every integer m spaces W_2^m and H_m are equivalent, i.e. they coincide as sets, and their norms are equivalent: there exist constants N_1, N_2 such that for every $u \in H_m$*

$$N_1|u|_{H_m} \le |u|_{W_2^m} \le N_2|u|_{H_m}.$$

From now on we identify spaces W_2^m and H_m, and denote both by W_2^m. We also preserve notations $|\cdot|_{W_2^m}$ for the norm, and $(\cdot\,,\cdot)_{W_2^m}$ for the scalar product. Below we list some properties of spaces W_2^m which follow from the properties of spaces H_m.

Proposition 4.2.2.

(i) For any integers m, n the system of spaces $(W_2^{m+n}, W_2^m, W_2^{m-n})$ forms a normal triple, and the mapping $\mathcal{J} : u \to (\cdot\,,u)$ generated by the scalar product in W_2^m (see Proposition 3.2.1) defines an isometric isomorphism between the spaces W_2^{m-n} and $(W_2^{m+n})^*$.

(ii) The space C_0^∞ is dense in W_2^m for every integer m in the topology of the latter space.

(iii) For all positive integer m, n, $\Lambda^n W_2^m = W_2^{m-n}$, and $(\Lambda^{2n}u, v)_{W_2^m} = (u, v)_{W_2^{m+n}}$ for all $u \in W_2^{m+2n}$, $v \in W_2^{m+n}$.

The next assertion is a fundamental result from the theory of Sobolev spaces.

Theorem 4.2.3 (Sobolev imbedding theorem). *The space W_2^m is normally imbedded into L_2. If for some non-negative integer n the inequality $2(m-n) > d$ holds then W_2^m is normally imbedded into the space C^n.*

The following result is very useful.

Proposition 4.2.4. *Let m be a positive integer.*

(i) There exist a constant N such that for all $v \in W_2^m$

$$\left|\frac{\partial}{\partial x_j} v\right|_{W_2^{m-1}} \leq N|v|_{W_2^m}.$$

(ii) For any $u, v \in W_2^1$

$$\int_{\mathbb{R}^d} u(x)\frac{\partial}{\partial x_j} v(x)dx = -\int_{\mathbb{R}^d} \left(\frac{\partial}{\partial x_j} u(x)\right) v(x)dx.$$

4.2.2 Cauchy Problem for Linear Equation of Second Order

On a given complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$ we consider a stochastic partial differential equation of the form

$$\begin{aligned} du(t,x) &= (D_p(a^{pq}(t,x)D_q u(t,x)) + f(t,x))dt \\ &\quad + (b_j^p(t,x)D_p u(t,x) + g_j(t,x))dW^j(t) \end{aligned} \tag{4.2.1}$$

on the interval $[0,T]$ (for some $T > 0$), with Cauchy condition

$$u(0,x) = u_0(x), \tag{4.2.2}$$

where a^{pq}, b_j^p are measurable in (t,ω,x) bounded real-valued functions on $[0,T] \times \Omega \times \mathbb{R}^d$; $f(t,\cdot)$, $g_j(t,\cdot)$ are stochastic fields with values in L_2, W_2^1, respectively, for all $p, q = 0, 1, \ldots, d$, $j = 1, 2, \ldots, d$; W is a d-dimensional Wiener process. We recall that repeated indices in monomials are summed over, i.e. in the right hand side of the equation (4.2.1) the first term is summed with respect to p, q, the second term is summed with respect to p, j.

We will use the following notations. For a separable Banach space $\mathbb{V}$ we denote by $C\mathbb{V}$ the class of $\mathcal{F}_t$-adapted continuous processes with values in $\mathbb{V}$. By $\mathcal{L}_p\mathbb{V}$ we denote the class of well-measurable processes with values in $\mathbb{V}$ for which almost surely

$$\int_0^T |u(t)|_{\mathbb{V}}^p dt < \infty,$$

where $|\cdot|_{\mathbb{V}}$ denotes the norm in $\mathbb{V}$. For two separable Banach spaces $\mathbb{U}$, $\mathbb{V}$ such that $\mathbb{V} \hookrightarrow \mathbb{U}$ we write $u \in C\mathbb{U} \cap \mathcal{L}_p\mathbb{V}$ if $u \in C\mathbb{U}$ and there exist process $\bar{u} \in \mathcal{L}_p\mathbb{V}$ such that $u = \bar{u}$ for almost every $(t, \omega) \in [0, T] \times \Omega$.

Definition 4.2.1. *We say that u is an L_2-solution of problem (4.2.1)-(4.2.2) on the interval $[0, T]$ if $u \in CL_2 \cap \mathcal{L}_2 W_2^1$, and for all $\varphi \in C_0^\infty$*

$$\begin{aligned}(u(t), \varphi)_{L_2} &= (u_0, \varphi)_{L_2} + \int_0^t \{(-1)^{\bar{p}}(a^{pq}(s) D_q u(s), D_p \varphi)_{L_2} + (f(s), \varphi)_{L_2}\} ds \\ &\quad + \int_0^t (b_j^p(s) D_p u(s) + g_j(s), \varphi)_{L_2} dW^j(s),\end{aligned}$$

for every $t \in [0, T]$ and almost every ω, where $\bar{p}$ assumes 0 for $p = 0$, and 1 otherwise, and the last integral is understood in the Itô sense.

4.3 The Main Result

Let W be an r-dimensional Wiener process, and $\{W_n\}_{n \in \mathbb{N}}$ a sequence of d-dimensional processes of bounded variation. Let us fix some positive number α. Suppose that the following holds.

Assumption 4.3.1. *For every $\kappa < \alpha$ and every positive δ*

$$\begin{aligned}&(1) \quad W - W_n = O(n^{-\kappa}), \\ &(2) \quad S_n = O(n^{-\kappa}), \\ &(3) \quad ||S_n|| = O(\ln^\delta n)\end{aligned}$$

on $[0, T]$, where S_n is an $r \times r$-dimensional process defined as follows,

$$S_n^{jl}(t) = \int_0^t (W^j(s) - W_n^j(s)) dW_n^l(s) - \frac{1}{2}\delta_{jl} t,$$

where δ_{jl} is the Kronecker's symbol which assumes 1 if $j = l$, and 0 otherwise.

Recall that given a function $F(t)$ expression $||F||(t)$ denotes its total variation over the interval $[0, t]$.

For every $n \geq 0$, $j = 1, 2, \ldots, r$, $p, q = 1, 2, \ldots, d$ let a_n^{pq}, b_{nj}^p be measurable in (t, ω, x) real-valued bounded functions on $[0, T] \times \Omega \times \mathbb{R}^d$. Let $f_n(t, \cdot)$ belong to $\mathcal{L}_2 L_2(\mathbb{R}^d)$, and $g_{nj}(t, \cdot)$ belong to $\mathcal{L}_2 W_2^1(\mathbb{R}^d)$ for $j = 1, 2, \ldots, d$.

Assumption 4.3.2. *For all $j = 1, 2, \ldots, r$, $p = 0, 1, \ldots, d$ there exist measurable in (t, ω, x) real-valued functions $b_j^{p(0)}, \ldots, b_j^{p(m)}$, $b_{nj}^{p(0)}, \ldots, b_{nj}^{p(m)}$, $n \in \mathbb{N}$, defined on $[0,T]\times\Omega\times\mathbb{R}^d$, and stochastic process $g_j^{(0)}, \ldots, g_j^{(m)}$, $g_{nj}^{(0)}, \ldots, g_{nj}^{(m)}$, $n \in \mathbb{N}$, belonging to $\mathcal{L}_2W_2^1(\mathbb{R}^d)$ such that for every $\varphi \in C_0^\infty(\mathbb{R}^d)$ the processes $(b_j^p(t), \varphi)$, $(b_{nj}^p(t), \varphi)$, $(g_j(t), \varphi)$, $(g_{nj}(t), \varphi)$ have on $[0,T]$ the following stochastic differentials,*

$$\begin{aligned} d(b_j^p(t), \varphi) &= (b_j^{p(0)}(t), \varphi)dt + (b_j^{p(l)}(t), \varphi)dW^l(t), \\ d(b_{nj}^p(t), \varphi) &= (b_{nj}^{p(0)}(t), \varphi)dt + (b_{nj}^{p(l)}(t), \varphi)dW_n^l(t), \end{aligned} \tag{4.3.1}$$

$$\begin{aligned} d(g_j(t), \varphi) &= (g_j^{(0)}(t), \varphi)dt + (g_j^{(l)}(t), \varphi)dW^l(t), \\ d(g_{nj}(t), \varphi) &= (g_{nj}^{(0)}(t), \varphi)dt + (g_{nj}^{(l)}(t), \varphi)dW_n^l(t). \end{aligned} \tag{4.3.2}$$

Moreover, for all $j, k = 1, 2, \ldots, r$, $p = 0, 1, \ldots, d$ there exist measurable in (t, ω, x) real-valued functions $b_j^{p(k0)}, \ldots, b_j^{p(km)}$ defined on $[0,T] \times \Omega \times \mathbb{R}^d$, and stochastic process $g_j^{(k0)}, \ldots, g_j^{(km)}$ belonging to $\mathcal{L}_2W_2^1(\mathbb{R}^d)$ such that for every $\varphi \in C_0^\infty(\mathbb{R}^d)$ the processes $(b_j^{p(k)}(t), \varphi)$, $(g_j^{(k)}(t), \varphi)$ have on $[0,T]$ the following stochastic differentials,

$$d(b_j^{p(k)}(t), \varphi) = (b_j^{p(k0)}(t), \varphi)dt + (b_j^{p(kl)}(t), \varphi)dW^l(t), \tag{4.3.3}$$

$$d(g_j^{(k)}(t), \varphi) = (g_j^{(k0)}(t), \varphi)dt + (g_j^{(kl)}(t), \varphi)dW^l(t). \tag{4.3.4}$$

We consider stochastic partial differential equation

$$\begin{aligned} u(t,x) &= \int_0^t (D_p(a^{pq}(s,x)D_q u(s,x)) + f(s,x))ds \\ &\quad + \int_0^t (b_j^p(s,x)D_p u(s,x) + g_j(s,x)) \circ dW^j(s), \end{aligned} \tag{4.3.5}$$

where $u_0(x) = u(0,x)$ is an $\mathcal{F}_t$-adapted random variable with values in $W_2^2(\mathbb{R}^d)$. Here we used the notation for the Stratonovich differential

$$\begin{aligned} &(b_j^p(t,x)D_p u(t,x) + g_j(t,x)) \circ dW^j(t) = \\ &\qquad = (b_j^p(t,x)D_p u(t,x) + g_j(t,x))dW^j(t) \\ &\qquad\quad + \frac{1}{2} b_j^p(t,x)D_p(b_j^q(t,x)D_q u(t,x) + g_j(t,x))dt \\ &\qquad\quad + \frac{1}{2}(b_j^{p(j)}(t,x)D_p u(t,x) + g_j^{(j)}(t,x))dt \end{aligned}$$

for fixed j. For every $n \in \mathbb{N}$ we consider partial differential equation

$$\begin{aligned} u_n(t,x) &= \int_0^t (D_p(a^{pq}(s,x)D_q u_n(s,x)) + f(s,x))ds \\ &\quad + (b_j^p(s,x)D_p u_n(s,x) + g_j(s,x))dW_n^j(s), \end{aligned} \tag{4.3.6}$$

where $u_{n0}(x) = u_n(0,x)$ is an $\mathcal{F}_t$-adapted random variable with values in $W_2^1(\mathbb{R}^d)$. We assume the following.

Assumption 4.3.3.

(i) The derivatives in x *of* a^{0q}, b_j^0 *up to the order* $m+4$, *of* a^{sq}, b_j^s *up to the order* $m+5$, *of* $b_j^{0(l)}$ *up to the order* $m+3$, *of* $b_j^{s(l)}$ *up to the order* $m+4$, *of* $b_j^{0(kl)}$ *up to the order* $m+1$, *of* $b_j^{s(kl)}$ *up to the order* $m+2$ *are bounded measurable in* (t,ω,x) *functions of* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^d$. *The derivatives in* x *of* a_n^{0q}, b_{nj}^0 *up to the order* m, *of* a_n^{sq}, b_{nj}^s *up to the order* $m+1$, *of* $b_{nj}^{0(l)}$ *up to the order* $\max\{m-1,0\}$, *of* $b_{nj}^{s(l)}$ *up to the order* m *are measurable in* (t,ω,x) *functions of* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^d$ *bounded uniformly in* n. *Above* $p,q=0,1,\dots,d$, $s=1,2,\dots,d$, $j,k=1,2,\dots,r$, $l=0,1,\dots,r$.

(ii) The random fields $f(t)$, $g_j(t)$, $g_j^{(l)}(t)$, $g_j^{(kl)}(t)$ *are well-measurable* $\mathcal{F}_t$*-adapted stochastic processes with values in* W_2^{m+4}, W_2^{m+5}, W_2^{m+4}, W_2^{m+2}, *respectively, such that*

$$|f(t)|_{W_2^{m+4}} \le K, \quad |g_j(t)|_{W_2^{m+5}} \le K, \quad |g_j^{(l)}(t)|_{W_2^{m+4}} \le K,$$
$$|g_j^{(kl)}(t)|_{W_2^{m+2}} \le K$$

for every $(t,\omega)\in[0,T]\times\Omega$ *and* $j,k=1,2,\dots,r$, $l=0,1,\dots,r$. *For every* $n\in\mathbb{N}$ *the random fields* $f_n(t)$, $g_{nj}(t)$, $g_{nj}^{(l)}(t)$ *are well-measurable* $\mathcal{F}_t$*-adapted stochastic processes with values in* W_2^m, W_2^{m+1}, W_2^m, *respectively, such that*

$$|f_n(t)|_{W_2^m} \le K, \quad |g_{nj}(t)|_{W_2^{m+1}} \le K, \quad |g_{nj}^{(l)}(t)|_{W_2^m} \le K$$

for every $(t,\omega)\in[0,T]\times\Omega$ *and* $j=1,2,\dots,r$, $l=0,1,\dots,r$. *The constant* K *above does not depend on* n.

(iii) The initial value u_0 *is an* $\mathcal{F}_0$*-measurable random variable with values in* W_2^{m+5}. *For every* $n\in\mathbb{N}$ *the initial value* u_{n0} *is an* $\mathcal{F}_0$*-measurable random variable with values in* W_2^{m+1}.

Assumption 4.3.4. *For all positive* $\kappa<\alpha$

(i) for all multi-indices γ, β *such that* $|\gamma|\le m$, $|\beta|\le m+1$

$$\begin{aligned} D^\gamma(a_n^{pq}(t,x)-a^{pq}(t,x)) &= O(n^{-\kappa}),\\ D^\beta(b_{nj}^p(t,x)-b_j^p(t,x)) &= O(n^{-\kappa}),\\ D^\gamma(b_{nj}^{p(l)}(t,x)-b_j^{p(l)}(t,x)) &= O(n^{-\kappa}) \end{aligned}$$

for every $x\in\mathbb{R}^d$ *and* $p,q=0,1,\dots,d$, $j,l=1,2,\dots,r$;

(ii)

$$\begin{aligned} f - f_n &= O(n^{-\kappa}) \text{ in } W_2^m, \\ g_j - g_{nj} &= O(n^{-\kappa}) \text{ in } W_2^{m+1}, \\ g_j^{(l)} - g_{nj}^{(l)} &= O(n^{-\kappa}) \text{ in } W_2^m \end{aligned}$$

for all $j, l = 1, 2, \ldots, r$;

(iii)

$$\xi - \xi_n = O(n^{-\kappa}) \text{ in } W_2^m.$$

Note that in the assumption above we do not require $l = 0$.

Assumption 4.3.5 (Strong ellipticity). *There exists a constant* $\lambda > 0$ *such that*

$$\sum_{p,q=1}^{d} a_n^{pq}(t, \omega, x) z_j z_l \geq \lambda |z|^2$$

for all $(t, \omega, x) \in [0, T] \times \Omega \times \mathbb{R}^d$ *and* $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$.

The existence of the solutions u, u_n under Assumptions 4.3.2, 4.3.3, 4.3.5 is well-known. It can be shown, for example, applying the technique used in [7] (see Propositions 4.12, 4.13, and the proof of Theorem 3.3). This is not the subject of the paper. Therefore, we concentrate on the estimation of the rate of convergence. We assume that there exist L_2-solutions u, u_n of the equations (4.3.5), (4.3.6), respectively, such that u is from the class $CW_2^{m+5} \cap \mathcal{L}_2 W_2^{m+6}$, u_n is from the class $CW_2^{m+1} \cap \mathcal{L}_2 W_2^{m+2}$.

Theorem 4.3.1. *Under the Assumptions 4.3.1-4.3.5 the sequence of solutions* u_n *of differential equations (4.3.6) converges almost surely to the solution* u *of "Stratonovich" stochastic differential equation (4.3.5). Moreover, for every* $\kappa < \alpha$

$$|u - u_n|_{W_2^m}^2 = O(n^{-\kappa}),$$

and

$$\int_0^T |u(s) - u_n(s)|_{W_2^{m+1}}^2 ds = O(n^{-\kappa}).$$

4.4 Preliminaries

In this section we list a set of technical lemmas which are borrowed from [7]. We will use them to prove Theorem 4.3.1. The proofs of all lemmas except Lemma 4.4.4 are omitted and can be found in [7]. Lemma 4.4.4 is a modification of Lemma 4.7 from [7]. Its proof is given.

We fix some non-negative integer m and consider the normal triple

$$W_2^{m+1} \hookrightarrow W_2^m \equiv W_2^{m*} \hookrightarrow W_2^{m-1}.$$

Lemma 4.4.1. *For any $p, q = 0, 1, \ldots, d$ let a^{pq} be a measurable function on $\mathbb{R}^d$. Suppose that for $s = 1, 2, \ldots, d$, $q = 0, 1, \ldots, d$ all derivatives of a^{0q} up to the order $\max(m-1, 0)$ and all derivatives of a^{sq} up to the order m are measurable bounded by a constant K. If for some constant $\lambda > 0$*

$$\sum_{j,l=1}^{d} a^{jl}(x) z_j z_l \;\geq\; \lambda |z|^2$$

for all $x \in \mathbb{R}^d$, $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, then

$$(D_p(a^{pq} D_q v), v)_{W_2^m} \;+\; \frac{\lambda}{2} |v|^2_{W_2^{m+1}} \;\leq\; L |v|^2_{W_2^m}$$

for every $v \in W_2^{m+1}$, where L is a constant depending only on K, m, λ and d.

Lemma 4.4.2. *Let a^{pq}, b^p be measurable bounded functions on $\mathbb{R}^d$ for $p, q = 0, 1, \ldots, d$.*

(i) Let us fix a non-negative integer α. If for every $s = 1, 2, \ldots, d$, $q = 0, 1, \ldots, d$ all derivatives of a^{0q}, b^0 up to the order $\max\{m + \alpha - 1, 0\}$ and all derivatives of a^{sq}, b^s up to the order $\max\{m + \alpha, 0\}$ are measurable bounded by a constant K functions, then for some constant L

$$|(D_p(a^{pq} D_q v), \varphi)_{W_2^m}| \;\leq\; L |v|_{W_2^{m+1+\alpha}} |\varphi|_{W_2^{m+1-\alpha}},$$

$$|(b^p D_p v, \varphi)_{W_2^m}| \;\leq\; L |v|_{W_2^{m+1+\alpha}} |\varphi|_{W_2^{m-\alpha}}$$

for every $v \in W_2^{m+1+\alpha}$, $\varphi \in C_0^\infty$.

(ii) If all derivatives of b^0 up to the order $\max\{m, 1\}$ and for every $s = 1, 2, \ldots, d$ the derivatives of b^s up to the order $\max\{m, 2\}$ are measurable bounded by a constant K functions, then for every $\alpha = 0, \pm 1, \pm 2$

$$|(b^p D_p v, \varphi)_{W_2^m}| \;\leq\; L |v|_{W_2^{m+\alpha}} |\varphi|_{W_2^{m+1-\alpha}}$$

for every $v \in W_2^{m+\alpha}$, $\varphi \in C_0^\infty$.

The constant L in the above statements depends only on K, m, d.

Lemma 4.4.3. *Let a^{pq}, b^p, c^p be bounded measurable functions on $\mathbb{R}^d$ for $p, q = 0, 1, \ldots, d$. Let us consider the differential operators $\mathcal{L}v = D_p(a^{pq} D_q v)$, $\mathcal{M}v = b^p D_p v$, $\mathcal{N}v = c^p D_p v$.*

(i) *If all derivatives of* c^0 *up to the order* m *and for* $s = 1, 2, \ldots, d$ *all derivatives of* c^s *up to the order* $\max(m, 1)$ *are measurable bounded by a constant* K *functions, then*

$$|(\mathcal{N}v, v)_{W_2^m}| \leq L|v|^2_{W_2^m}$$

for all $v \in W_2^{m+1}$, *where* L *is a constant.*

(ii) *If all derivatives of* b^0, c^0 *up to the order* $m+1$ *and for* $s = 1, 2, \ldots, d$ *all derivatives of* b^s, c^s *up to the order* $\max(m + 1, 2)$ *are measurable bounded by a constant* K *functions, then*

$$|(\mathcal{M}v, \mathcal{N}v)_{W_2^m} + (v, \mathcal{M}\mathcal{N}v)_{W_2^m}| \leq L|v|^2_{W_2^{m+1}}$$

for all $v \in W_2^{m+1}$, *where* L *is a constant.*

(iii) *If all derivatives of* b^0 *up to the order* $m + 1$, *for* $s = 1, 2, \ldots, d$ *all derivatives of* b^s *up to the order* $\max(m+1, 1)$ *for every* $p, q = 0, 1, \ldots, d$ *all derivatives of* a^{pq} *up to the order* m *are measurable bounded by a constant* K *functions, then for some constant* L

$$|(\mathcal{L}v, \mathcal{M}v)_{W_2^m} + (v, \mathcal{M}\mathcal{L}v)_{W_2^m}| \leq L|v|^2_{W_2^{m+1}}$$

for all $v \in W_2^{m+2}$.

The constant L *in the above statements depends only on* K, m, d.

Lemma 4.4.4. *Let* $c_n = c_n(t, \omega, x)$ *be a* $\mathcal{B}([0, T]) \times \mathcal{F} \times \mathcal{B}(\mathbb{R}^d)$*-measurable real function on* $[0, T] \times \Omega \times \mathbb{R}^d$ *for every* $n = 1, 2, \ldots$, *such that its derivatives in* x *up to the order* r *are measurable functions of* (t, ω, x), *bounded by a constant* K, *where* r *is a non-negative integer. Suppose that for some* $\kappa > 0$

$$D^\gamma c_n(t, x) = O(n^{-\kappa})$$

uniformly in $x \in \mathbb{R}^d$ *and all multi-indices* γ, *such that* $|\gamma| \leq r$. *Then*

(i) *if* $r = m$, *then*

$$D_p(c_n(t)D_q v) = O(n^{-\kappa}) \text{ in } W_2^{m-1},$$

$$c_n(t)D_q v = O(n^{-\kappa}) \text{ in } W_2^m$$

for every $v \in W_2^{m+1}$;

(ii) *if* $r = \max(m - 1, 0)$, *then*

$$c_n(t)D_q v = O(n^{-\kappa}) \text{ in } W_2^{m-1}$$

for every $v \in W_2^m$;

(iii) if $r = m+1$, then

$$c_n(t)D_q v = O(n^{-\kappa}) \text{ in } W_2^{m+1}$$

for every $v \in W_2^{m+2}$.

In order to prove this lemma we will need the following.

Lemma 4.4.5. *For any $m \in \mathbb{Z}$ and every $v \in W_2^{m-1}$*

$$|v|_{W_2^{m-1}} = \sup_{|\varphi|_{W_2^{m+1}} \leq 1} (v, \varphi)_{W_2^m},$$

where $(\cdot, \cdot)_{W_2^m}$ denotes the scalar product in W_2^m and the duality between the spaces W_2^{m+1}, W_2^{m-1}.

Proof of Lemma 4.4.5. Let us introduce a self-adjoined positive definite operator $\Lambda = (I - \Delta)^{1/2}$, where Δ is the Laplace operator, I is the identity operator. Then (see Section 4.2.1)

$$|v|^2_{W_2^{m-1}} = (\Lambda^{m-1}v, \Lambda^{m-1}v).$$

Dividing both sides by $|v|_{W_2^{m-1}}$ we get

$$|v|_{W_2^{m-1}} = (\Lambda^{m-1}v, \Lambda^{m-1}\bar{v}), \qquad \bar{v} = \frac{v}{|v|_{W_2^{m-1}}},$$

and then, since the operator Λ is self-adjoint,

$$|v|_{W_2^{m-1}} = (\Lambda^m v, \Lambda^m \varphi) = (v, \varphi)_{W_2^m}, \qquad \varphi = \Lambda^{-2}\bar{v}. \tag{4.4.1}$$

Again using the representation $|\cdot|_{W_2^{m+1}} = |\Lambda^{m+1} \cdot|_{L_2}$ it is easy to show that $|\varphi|_{W_2^{m+1}} = 1$. Now, taking the supremum in the right hand side of (4.4.1) we get

$$|v|_{W_2^{m-1}} \leq \sup_{|\varphi|_{W_2^{m+1}} \leq 1} (v, \varphi)_{W_2^m}.$$

On the other side for any $\varphi \in W_2^{m+1}$

$$(v, \varphi)_{W_2^m} = (\Lambda^m v, \Lambda^m \varphi)_{L_2} \leq |v|_{W_2^{m-1}} |\varphi|_{W_2^{m+1}},$$

which implies

$$\sup_{|\varphi|_{W_2^{m+1}} \leq 1} (v, \varphi)_{W_2^m} \leq |v|_{W_2^{m-1}}.$$

□

Proof of Lemma 4.4.4. To prove the first assertion we produce the following estimations. Using Lemma 4.4.5, Proposition 4.2.4 and Cauchy-Bunyakovskii-Schwartz inequality, for any $v \in W_2^{m+1}$

$$\begin{aligned}
|D_p(c_n(t)D_qv)|^2_{W_2^{m-1}} &= \sup_{|\varphi|_{W_2^{m+1}}\leq 1} |(D_p(c_n(t)D_qv), \varphi)_{W_2^m}|^2 \\
&= \sup_{|\varphi|_{W_2^{m+1}}\leq 1} |(c_n(t)D_qv, D_p\varphi)_{W_2^m}|^2 \\
&\leq |c_n(t)D_qv|^2_{W_2^m} \\
&= \sum_{|\gamma|\leq m} \int_{\mathbb{R}^d} D^\gamma |c_n(t,x)D_qv(x)|^2 dx,
\end{aligned}$$

which immediately gives the result due to the assumptions of the lemma. The rest of the lemma can be proved similarly. □

4.5 Proof of Theorem 4.3.1

We reduce the proof of the theorem to verification of the assumptions of Theorem 3.3.1. In connection with the notations in Theorem 3.3.1 we denote $\mathbb{H}_0 = W_2^m$, $\mathbb{H}_\beta = W_2^{m+\beta}$ for any positive or negative integer β, and denote $|\cdot|_\beta$ the norm in the space $\mathbb{H}_\beta$, and $|\cdot|$, $(\cdot\,,\cdot)$ the norm and the scalar product in the Hilbert space $\mathbb{H}_0$.

We consider a normal triple

$$\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \equiv \mathbb{H}_0^* \hookrightarrow \mathbb{H}_1^*,$$

and, moreover, the system of embedded spaces

$$\mathbb{H}_6 \hookrightarrow \mathbb{H}_5 \hookrightarrow \mathbb{H}_4 \hookrightarrow \mathbb{H}_3 \hookrightarrow \mathbb{H}_2 \hookrightarrow \mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \equiv \mathbb{H}_0^* \hookrightarrow \mathbb{H}_1^* \hookrightarrow \mathbb{H}_2^* \hookrightarrow \mathbb{H}_3^*.$$

We recall that spaces $\mathbb{H}_0$ and $\mathbb{H}_0^*$, as well as $\mathbb{H}_{-\beta}$ and $\mathbb{H}_\beta^*$ can be naturally identified using scalar multiplication $(\cdot\,,\cdot)$ (see section 3.2.1). We use the notation (u,v) for the duality between $\mathbb{H}_\beta$ and $\mathbb{H}_\beta^*$ where one of the elements u, v belongs to $\mathbb{H}_\beta$ and the other to $\mathbb{H}_\beta^*$.

Let us introduce for all $v \in W_2^{m+1}$, $(t,\omega) \in [0,T] \times \Omega$ operators

$$\begin{aligned}
\mathcal{L}v &= D_p(a^{pq}D_qv), \quad \mathcal{M}^jv = b_j^pD_pv, \quad \mathcal{M}^{j(l)}v = b_j^{p(l)}D_pv, \\
&\qquad \mathcal{M}^{j(kl)}v = b_j^{p(kl)}D_pv, \\
\mathcal{L}_nv &= D_p(a_n^{pq}D_qv), \quad \mathcal{M}_n^jv = b_{nj}^pD_pv, \quad \mathcal{M}_n^{j(l)}v = b_{nj}^{p(l)}D_pv,
\end{aligned}$$

$j,k=1,2,\dots,r$, $l=0,1,\dots,r$, and consider equation

$$\begin{aligned}(u(t),\varphi) &= (u_0,\varphi) + \int_0^t (\mathcal{L}u(s)+f(s),\varphi)ds \\ &+ \frac{1}{2}\int_0^t (\mathcal{M}^j\mathcal{M}^j u(s)+\mathcal{M}^j g_j(s),\varphi)ds \\ &+ \frac{1}{2}\int_0^t (\mathcal{M}^{j(j)}u(s)+g_j^{(j)}(s),\varphi)ds \\ &+ \int_0^t (\mathcal{M}^j u(s)+g_j(s),\varphi)dW^j(s),\end{aligned} \tag{4.5.1}$$

and for every $n=1,2,\dots$ equation

$$\begin{aligned}(u_n(t),\varphi) &= (u_{n0},\varphi) + \int_0^t (\mathcal{L}_n u_n(s)+f_n(s),\varphi)ds \\ &+ \int_0^t (\mathcal{M}_n^j u_n(s)+g_{nj}(s),\varphi)dW_n^j(s),\end{aligned} \tag{4.5.2}$$

where $\varphi\in C_0^\infty$. By a solution u of equation (4.5.1) and a solution u_n of equation (4.5.2) we mean functions from $CW_2^m\cap\mathcal{L}_2W_2^{m+1}$ for which the equations (4.5.1), (4.5.2) hold respectively almost surely for every $\varphi\in C_0^\infty(\mathbb{R}^d)$ and $t\in[0,T]$.

From Assumption 4.3.3 using statement (i) of Lemma 4.4.2 it follows that for all $j,k=1,2,\dots,r$, $l=0,1,\dots,r$, every $(t,\omega)\in[0,T]\times\Omega$

$$\begin{aligned}(\mathcal{L}v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^{m+1}}, \\ (\mathcal{M}^j v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^m}, \\ (\mathcal{M}^{j(l)} v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^m}, \\ (\mathcal{M}^{j(kl)} v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^m},\end{aligned} \tag{4.5.3}$$

and for all $n\in\mathbb{N}$

$$\begin{aligned}(\mathcal{L}_n v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^{m+1}}, \\ (\mathcal{M}_n^j v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^m}, \\ (\mathcal{M}_n^{j(l)} v,\varphi)_{W_2^m} &\le L|v|_{W_2^{m+1}}|\varphi|_{W_2^m}\end{aligned} \tag{4.5.4}$$

for every $v\in W_2^{m+1}$, $\varphi\in C_0^\infty$. Then for every $j,k=1,2,\dots,r$, $l=0,1,\dots,r$ by

$$\begin{aligned}(A(t)v,\varphi) &= (\mathcal{L}v,\varphi)_{W_2^m}, \\ (B^j(t)v,\varphi) &= (\mathcal{M}^j v,\varphi)_{W_2^m}, \\ (B^{j(l)}(t)v,\varphi) &= (\mathcal{M}^{j(l)} v,\varphi)_{W_2^m} \\ (B^{j(kl)}(t)v,\varphi) &= (\mathcal{M}^{j(kl)} v,\varphi)_{W_2^m}\end{aligned}$$

we define bounded linear operators $A(t) : \mathbb{H}_1 \to \mathbb{H}_{-1}$, $B^j(t), B^{j(l)}(t), B^{j(kl)}(t)$: $\mathbb{H}_1 \to \mathbb{H}_0$, and by

$$\begin{aligned}
(A_n(t)v, \varphi) &= (\mathcal{L}_n v, \varphi)_{W_2^m}, \\
(B_n^j(t)v, \varphi) &= (\mathcal{M}_n^j v, \varphi)_{W_2^m}, \\
(B_n^{j(l)}(t)v, \varphi) &= (\mathcal{M}_n^{j(l)} v, \varphi)_{W_2^m}
\end{aligned}$$

we define bounded uniformly in n linear operators $A_n(t) : \mathbb{H}_1 \to \mathbb{H}_{-1}$, $B_n^j(t), B_n^{j(l)}(t)$: $\mathbb{H}_1 \to \mathbb{H}_0$.

Moreover, by Lemma 4.4.2 for all $j, l = 1, 2, \ldots, r$ and every $(t, \omega) \in [0, T] \times \Omega$ we have,

$$(\mathcal{M}^j(\mathcal{M}^l v), \varphi)_{W_2^m} \leq L|v|_{W_2^{m+1}} |\varphi|_{W_2^{m+1}}$$

for all $v \in W_2^{m+1}$, $\varphi \in C_0^\infty$. Again, by

$$(B^{jl}(t)v, \varphi) = (\mathcal{M}^j(\mathcal{M}^l v), \varphi)_{W_2^m}$$

we define bounded linear operators $B^{jl} : \mathbb{H}_1 \to \mathbb{H}_{-1}$. Statement (ii) of Lemma 4.4.2 gives for $j = 1, 2, \ldots, r$ and $(t, \omega) \in [0, T] \times \Omega$

$$(\mathcal{M}^j v, \varphi)_{W_2^m} \leq L|v|_{W_2^m} |\varphi|_{W_2^{m+1}}$$

for all $v \in W_2^m$, $\varphi \in C_0^\infty$. This allows us to extend operator $B^j : \mathbb{H}_1 \to \mathbb{H}_0$ to a bounded operator from $\mathbb{H}_0$ to $\mathbb{H}_{-1}$, and for all (t, ω) operator $B^{jl} : \mathbb{H}_1 \to \mathbb{H}_{-1}$ is a superposition of $B^l : \mathbb{H}_1 \to \mathbb{H}_0$ and $B^j : \mathbb{H}_0 \to \mathbb{H}_{-1}$.

From the fact that $(w, (I - \Delta)^m \varphi)_{L_2} = (w, \varphi)_{W_2^m}$ for all $v, \varphi \in C_0^\infty$ it follows from (4.3.1) that substituting $(D_p v)(I - \Delta)^m \varphi$ in place of φ we get for $j = 1, 2, \ldots, r$, $p = 0, 1, \ldots, d$

$$d(b_j^p(t) D_p v, \varphi)_{W_2^m} = (b_j^{p(0)}(t) D_p v, \varphi)_{W_2^m} dt + (b_j^{p(l)}(t) D_p v, \varphi)_{W_2^m} dW^l$$

and for every $n = 1, 2, \ldots$

$$d(b_{nj}^p(t) D_p v, \varphi)_{W_2^m} = (b_{nj}^{p(0)}(t) D_p v, \varphi)_{W_2^m} dt + (b_{nj}^{p(l)}(t) D_p v, \varphi)_{W_2^m} dW_n^l,$$

which can be written as

$$d(B^j(t)v, \varphi) = (B^{j(0)}(t)v, \varphi)dt + (B^{j(l)}(t)v, \varphi)dW^l$$

and

$$d(B_n^j(t)v, \varphi) = (B_n^{j(0)}(t)v, \varphi)dt + (B_n^{j(l)}(t)v, \varphi)dW_n^l$$

for every $v, \varphi \in C_0^\infty$, and hence by (4.5.3), (4.5.4) for any $v, \varphi \in \mathbb{H}_1$. Similarly, substituting $(I - \Delta)^m \varphi$ in place of φ we get

$$d(g_j(t), \varphi) = (g_j^{(0)}(t), \varphi)dt + (g_j^{(l)}(t), \varphi)dW^l$$

and

$$d(g_{nj}(t),\varphi) \;=\; (g_{nj}^{(0)}(t),\varphi)dt \;+\; (g_{nj}^{(l)}(t),\varphi)dW_n^l$$

for every $\varphi \in C_0^\infty$, and hence for every $\varphi \in \mathbb{H}_1$. Using the same sort of arguments we can obtain from (4.3.3), for $j,k = 1,2,\dots,r$, $p = 0,1,\dots,d$

$$d(b_j^{p(k)}(t)D_p v,\varphi)_{W_2^m} \;=\; (b_j^{p(k0)}(t)D_p v,\varphi)_{W_2^m}dt \;+\; (b_j^{p(kl)}(t)D_p v,\varphi)_{W_2^m}dW^l$$

which can be written in the form

$$d(B^{j(k)}(t)v,\varphi) \;=\; (B^{j(k0)}(t)v,\varphi)dt \;+\; (B^{j(kl)}(t)v,\varphi)dW^l,$$

and from (4.3.4)

$$d(g_j^{(k)}(t),\varphi) \;=\; (g_j^{(k0)}(t),\varphi)dt \;+\; (g_j^{(kl)}(t),\varphi)dW^l$$

for every $v,\varphi \in \mathbb{H}_1$.

We rewrite equations (4.5.1), (4.5.2) in the form

$$\begin{aligned} u(t) \;&=\; \xi + \int_0^t (A(s)u(s) + f(s))ds \;+\; \int_0^t (B^j(t)u(t) + g_j(t))dW^j(t) \\ &\quad + \frac{1}{2}\int_0^t B^j(t)(B^j(t)u(t) + g_j(t))dt \;+\; \frac{1}{2}\int_0^t (B^{j(j)}(t)u(t) + g_j^{(j)}(t))dt. \end{aligned}$$

and

$$u_n(t) \;=\; \xi_n \;+\; \int_0^t (A_n(s)u_n(s) + f_n(s))ds \;+\; \int_0^t (B_n^j(s)u_n(s) + g_{nj}(s))dW_n^j(s)$$

which are considered in the triple $\mathbb{H}_1 \hookrightarrow \mathbb{H}_0 \equiv \mathbb{H}_0^* \hookrightarrow \mathbb{H}_1^*$. We can see that these equation are of the same type as those in Theorem 3.3.1. Therefore, it suffices to verify Assumptions 3.3.3 - 3.3.5 of Theorem 3.3.1.

By Lemma 4.4.2 for all $j = 1,2,\dots,r$, $l = 0,1,\dots,r$ for some constant L

$$\begin{aligned} |(A(t)v,\varphi)| \;&\le\; L|v|_k|\varphi|_{2-k}, \quad k = 0,1,\dots,6, \\ |(B^j(t)v,\varphi)| \;&\le\; L|v|_k|\varphi|_{1-k}, \quad k = -2,-1,\dots,6, \\ |(B^{j(l)}(t)v,\varphi)| \;&\le\; L|v|_k|\varphi|_{1-k}, \quad k = -2,-1,\dots,5, \\ |(B^{j(kl)}(t)v,\varphi)| \;&\le\; L|v|_3|\varphi|_{-2} \end{aligned}$$

for every $(t,\omega) \in [0,T]\times\Omega$ and $v,\varphi \in C_0^\infty$. Similarly, for all $n \in \mathbb{N}$, $j = 1,2,\dots,d$, $l = 0,1,\dots,r$ for some constant L_n

$$\begin{aligned} |(A_n(t)v,\varphi)| \;&\le\; L_n|v|_k|\varphi|_{2-k}, \quad k = 0,1,2, \\ |(B_n^j(t)v,\varphi)| \;&\le\; L_n|v|_k|\varphi|_{1-k}, \quad k = -1,0,1, \\ |(B_n^{j(l)}(t)v,\varphi)| \;&\le\; L_n|v|_k|\varphi|_{1-k}, \quad k = -1,0 \end{aligned}$$

for every $(t, \omega) \in [0, T] \times \Omega$ and $v, \varphi \in C_0^\infty$. Therefore, the requirements of Assumption 3.3.3 are satisfied.

Next, we verify requirements of Assumption 3.3.4. Using Lemma 4.4.1 we get for every $n \in \mathbb{N}$

$$(v, A_n v) + \frac{\lambda}{2}|v|_1^2 \leq L|v|_0^2$$

for every $(t, \omega) \in [0, T] \times \Omega$ and $v \in W_2^{m+1}$. Here λ is a constant that comes from Assumption 4.3.5. Using Lemma 4.4.3 we get for every $n \in \mathbb{N}$

$$\begin{aligned} |(B_n^j v, B_n^l v) + (v, B_n^j B_n^l v)| &\leq L|v|_0^2, \\ |(v, B_n^j v)| &\leq L|v|_0^2, \\ |(v, B_n^{j(l)} v)| &\leq L|v|_0^2, \end{aligned}$$

$$|(A_n w, B_n^j w) + (w, B_n^j A_n w)| \leq L|w|_1^2$$

for every $j, l = 1, 2, \ldots, r$, $(t, \omega) \in [0, T] \times \Omega$ and $v \in W_2^{m+1}$, $w \in W_2^{m+2}$. In inequalities above constant L does not depend on n.

Finally, Assumption 3.3.5 holds by Lemma 4.4.4. This completes the proof.

□

Chapter 5

Filtering Problem

This chapter is devoted to the application of the result of Chapter 4 to the filtering problem.

5.1 Introduction

The problem of non-linear filtering can be described as follows. Assume that $(x, y) = \{(x(t), y(t), t \geq 0)\}$ are two diffusion processes with values in $\mathbb{R}^d$, $\mathbb{R}^r$, respectively. The process of interest, the unobservable *signal process* x, represents the state of the system at time t. The observable component, the *observation process* y, represents the measured output of the system at time t. The problem of the estimation of the unobservable signal $x(t)$ or a function of $x(t)$ on the basis of the observed paths of $y(s)$ for $s \leq t$ is referred to as a filtering problem.

This model arises in many technical problems. For example, the process $y(t)$ describes the coordinates of a moving object computed on the basis of the radar measurements, $N(t)$ represents the error of the measurements, and $x(t)$ is the true position of the object at time t.

Here we have a simplified model. Suppose that x is a solution of ordinary differential equation $dx(t)/dt = H(x(t))$, $x(0) = x_0$. Then the observation y is a solution of equation

$$y(t) = \int_0^t H(x(s))ds + N(t).$$

In various situations the evolution equation of the signal process x includes random perturbation. In these cases x can be described by the equation

$$x(t) = \int_0^t h(x(s))ds + \int_0^t \sigma(x(s))dN(s).$$

Thus, the goal of the filtering problem is to filter out the noise N from the observation process y, and to find the best estimate for the signal process x given the measurements up to time t. More precisely, given a measurable function $f = f(x)$ with $E(f^2(x(t))) < \infty$ find the best mean square estimate $\hat{f}(t)$ of $f(x(t))$ given the trajectories $y(s)$, $s \leq t$. If the noise N is a Wiener process then it is known that under certain regularity assumptions

$$\hat{f}(t) = \frac{\int_{\mathbb{R}} f(x)\varphi(t,x)dx}{\int_{\mathbb{R}} \varphi(t,x)dx},$$

where φ is a random process, called *unnormalized filtering density.* The problem of estimating $f(x(t))$ is thus reduced to the problem of computing the density φ. It was also shown that φ satisfies a linear stochastic differential equation,

$$\begin{aligned} d\varphi(t,x) &= \left(\frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(x)\varphi(t,x)) - \frac{\partial}{\partial x}(h(x)\varphi(t,x))\right) dt \\ &\quad + \left(H(x)\varphi(t,x) - \frac{\partial}{\partial x}(\sigma(x)\varphi(t,x))\right) dy(t), \end{aligned}$$

called the Zakai equation (see [19]). The *normalized filtering density*

$$p(t,x) = \varphi(t,x)\left(\int_{\mathbb{R}} \varphi(t,x)dx\right)^{-1}$$

was also studied. It is known that p satisfies a measure valued stochastic differential equation called the Fujisaki-Kallianpur-Kunita equation with a disadvantage of being nonlinear (see [3]).

In the paper we study a more general situation. The processes x, y are considered to be multidimensional and functions h, H, σ depend on t, x, y. Our interest is motivated by the following. In practice the "real observations" have bounded variation. This is the result of the error in measurement of process y. We thus get a sequence $\{y_n\}_{n\in\mathbb{N}}$ of processes of bounded variation which approximates observation y. Using "real observations" y_n instead of y, we solve the approximation of the Zakai equation. We compute φ_n, and hence obtain density p_n by the formula

$$p_n(t,x) = \varphi_n(t,x)\left(\int_{\mathbb{R}} \varphi_n(t,x)dx\right)^{-1}.$$

It is known that under some general conditions the convergence $y_n \to y$ implies the convergence of unnormalized densities $\varphi_n \to \varphi$, as well as normalized densities $p_n \to p$. In Theorem 5.2.2 we study the rate of this convergences given the rate of convergence $y_n \to y$.

5.2 The Main Result

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space, and let us consider a $(d+r)$-dimensional diffusion process $(x,y) = \{(x(t),y(t)),\ t \in [0,T]\}$ defined by the system of stochastic differential equations

$$\begin{aligned} dx(t) &= h(t,x(t),y(t))dt + \sigma(t,x(t),y(t))dV(t) \\ &\quad + \rho(t,x(t),y(t))dW(t), \\ dy(t) &= H(t,x(t),y(t))dt + dW(t), \end{aligned} \tag{5.2.1}$$

$t \in (0,T]$, with the initial conditions

$$x(0) = \xi, \quad y(0) = \eta. \tag{5.2.2}$$

Here $h(t,x,y)$, $H(t,x,y)$, $\sigma(t,x,y)$, $\rho(t,x,y)$ are matrices of the size $d \times 1$, $r \times 1$, $d \times r_0$, $d \times r$, respectively (for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^r$), $(V(t), W(t))$ is a standard $(r_0 + r)$-dimensional Wiener process independent of the $\mathcal{F}_0$-measurabe random variables ξ, η with values in $\mathbb{R}^d$, $\mathbb{R}^r$, respectively. In the paper we assume that the model is non-degenerate, i.e. there exist $\varepsilon > 0$ such that

$$\rho\rho^*(t,x,y) \ge \varepsilon I$$

for all $(t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^r$. Above I is the identity operator.

Let us define $h^{pq} = h^{pq}(t,x,y)$ by

$$h^{pq} = \frac{1}{2}(\rho\rho^*)^{pq} + \frac{1}{2}(\sigma\sigma^*)^{pq}$$

for every $p, q = 1, 2, \ldots, d$, and define operators

$$\begin{aligned} \hat{\mathcal{L}}(t,x)u &= -\frac{\partial}{\partial x_p}(h^p(t,x,y(t))u) + \frac{\partial^2}{\partial x_p \partial x_q}(h^{pq}(t,x,y(t))u), \\ \hat{\mathcal{M}}^j(t,x)u &= H^j(t,x,y(t))u - \frac{\partial}{\partial x_p}(\sigma^{pj}(t,x,y(t))u), \end{aligned}$$

$j = 1, 2, \ldots, r$. Let us consider the so called Zakai equation

$$d\varphi(t,x) = \hat{\mathcal{L}}\varphi(t,x)dt + \hat{\mathcal{M}}^j\varphi(t,x)dy^j(t) \tag{5.2.3}$$

with the initial condition

$$\varphi(0,x) = p_0(x). \tag{5.2.4}$$

The following theorem is a fundamental result.

Theorem 5.2.1. *Let $l \ge 0$ be an integer. Assume that h^{pq} have uniformly bounded derivatives in x up to the order $l+2$, h and ρ have uniformly bounded derivatives in x up to the order $l+1$ and H have uniformly bounded derivatives in x up to the order l. Suppose that the conditional distribution of ξ given η has a density p_0 (with respect to Lebesgue measure), which belongs to W_2^l. Then the* *conditional density* $p(t,x) := \mathbf{P}\{t,dx\}/dx$ *(of* $x(t)$ *given* $\{y(s),\ 0 \le s \le t\}$*) exists and*

$$p(t,x) \;=\; \varphi(t,x)/(\varphi(t),1)_{L_2},$$

where $\varphi = \varphi(t,x)$ *is the unique solution of Zakai equation (5.2.3) with the initial condition (5.2.4). Moreover,* $\varphi \in CW_2^l \cap \mathcal{L}_2 W_2^{l+1}$ *and* $(\varphi(t),1)_{L_2}$ *is a positive and continuous function on* $[0,T]$.

Let us approximate observation process y by a sequence $\{y_n\}_{n\in\mathbb{N}}$ of r-dimensional continuous on $[0,T]$ stochastic processes of bounded variation. Similarly as above for every $n \in \mathbb{N}$ we define operators

$$\begin{aligned}
\hat{\mathcal{L}}_n(t,x)u &= -\frac{\partial}{\partial x_p}(h^p(t,x,y_n(t))u) + \frac{\partial^2}{\partial x_p \partial x_q}(h^{pq}(t,x,y_n(t))u),\\
\hat{\mathcal{M}}^j_n(t,x)u &= H^j(t,x,y_n(t))u - \frac{\partial}{\partial x_p}(\rho^{pj}(t,x,y_n(t))u),\\
\hat{\mathcal{M}}^{j(l)}_n(t,x)u &= \frac{\partial}{\partial y_l}H^j(t,x,y_n(t))u - \frac{\partial}{\partial x_p}(\frac{\partial}{\partial y_l}\rho^{pj}(t,x,y_n(t))u)
\end{aligned}$$

for $j,l = 1,2,\dots,r$. As an approximation for the problem (5.2.3)-(5.2.4) we consider the partial differential equation

$$\begin{aligned}
d\varphi_n(t,x) &= \hat{\mathcal{L}}_n\varphi_n(t,x)dt + \hat{\mathcal{M}}^j_n\varphi_n(t,x)dy^j_n(t)\\
&\quad - \frac{1}{2}\hat{\mathcal{M}}^j_n\hat{\mathcal{M}}^j_n\varphi_n(t,x)dt - \frac{1}{2}\hat{\mathcal{M}}^{j(j)}_n\varphi_n(t,x)dt
\end{aligned} \tag{5.2.5}$$

with the initial condition

$$\varphi_n(0,x) \;=\; p_0(x). \tag{5.2.6}$$

Suppose the following holds.

Assumption 5.2.1. *For every* $\kappa < \alpha$ *and every positive* δ

$$\begin{aligned}
&(1) \quad y - y_n = O(n^{-\kappa}),\\
&(2) \quad S_n = O(n^{-\kappa}),\\
&(3) \quad \|S_n\| = O(\ln^\delta n),
\end{aligned}$$

where S_n *is an* $r \times r$*-dimensional process defined as follows,*

$$S^{jl}_n(t) \;=\; \int_0^t (y^j(s) - y^j_n(s))dy^l_n(s) - \frac{1}{2}\delta_{jl}t,$$

where δ_{jl} *is the Kronecker's symbol which assumes* 1 *if* $j = l$*, and* 0 *otherwise.*

Assumption 5.2.2. *The derivatives in* x *of* h^p*,* H^j *up to the order* $m+5$*, of* $(\partial/\partial y_l)H^j$ *up to the order* $m+4$*, of* $(\partial/\partial t)H^j$*,* $(\partial^2/\partial y_l \partial y_h)H^j$ *up to the order*

$m+3$, of $(\partial^2/\partial t\partial y_l)H^j$, $(\partial^3/\partial y_l\partial y_h\partial y_k)H^j$ up to the order $m+1$, of σ^{pi}, ρ^{pj} up to the order $m+6$, of $(\partial/\partial y_l)\sigma^{pi}$, $(\partial/\partial y_l)\rho^{pj}$ up to the order $m+5$, of $(\partial/\partial t)\rho^{pj}$, $(\partial^2/\partial y_l\partial y_h)\rho^{pj}$ up to the order $m+4$, of $(\partial^2/\partial t\partial y_l)\rho^j$, $(\partial^3/\partial y_l\partial y_h\partial y_k)\rho^j$ up to the order $m+2$, are bounded measurable in (t,x,y) functions on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^r$. Above $p=1,2,\dots,d$, $j,l,h,k=1,2,\dots,r$, $i=1,2,\dots,r_0$.

Assumption 5.2.3. *Random variables ξ, η are almost surely finite. The conditional distribution of ξ given η is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ for almost every $\omega\in\Omega$. Its density $p_0=p_0(\cdot,\omega)$ belongs to W_2^{m+5}.*

Note, that under Assumptions 5.2.2, 5.2.3 there exist an L_2-solution φ, φ_n of problems (5.2.3)-(5.2.4), (5.2.5)-(5.2.6), respectively, such that φ belongs to $CW_2^{m+5}\cap\mathcal{L}_2W_2^{m+6}$, φ_n belongs to $CW_2^{m+1}\cap\mathcal{L}_2W_2^{m+2}$. As a corollary of Theorem 4.3.1 we get the following.

Theorem 5.2.2. *Under Assumptions 5.2.1-5.2.3 the sequence of solutions φ_n of the problem (5.2.5)-(5.2.6) converges almost surely to the solution φ of the problem (5.2.3)-(5.2.4),*

$$|\varphi-\varphi_n|^2_{W_2^m} \;=\; O(n^{-\kappa}),$$

and

$$\int_0^T|\varphi(s)-\varphi_n(s)|^2_{W_2^{m+1}}ds \;=\; O(n^{-\kappa})$$

for every $\kappa<\alpha$. Moreover, the sequence of densities p_n converges almost surely to the density p,

$$|p-p_n|^2_{W_2^m} \;=\; O(n^{-\kappa}),$$

and

$$\int_0^T|p(s)-p_n(s)|^2_{W_2^{m+1}}ds \;=\; O(n^{-\kappa})$$

for every $\kappa<\alpha$.

5.3 Proof of Theorem 5.2.2

5.3.1 Convergence $\varphi_n\to\varphi$

In order to define operators $\mathcal{M}^{j(l)}(t,x)$, $\mathcal{M}_n^{j(l)}(t,x)$, $\mathcal{M}^{j(kl)}(t,x)$, $\mathcal{M}_n^{j(kl)}(t,x)$ we rewrite the definition for $\mathcal{M}^j(t,x)$, $\mathcal{M}_n^j(t,x)$. For every $j=1,2,\dots,r$, $p=1,2,\dots,d$ set

$$\beta_j^0(t,x,y) \;=\; H^j(t,x,y) \;-\; \frac{\partial}{\partial x_p}\rho^{pj}(t,x,y), \qquad \beta_j^p(t,x,y) \;=\; -\rho^{pj}(t,x,y).$$

Define

$$b_j^p(t,x) \;=\; \beta_j^p(t,x,y(t)), \qquad b_{nj}^p(t,x) \;=\; \beta_{nj}^p(t,x,y_n(t)),$$

$p = 0,1,\dots,d$, and then in connection to the notations introduced in chapter 4

$$\hat{\mathcal{M}}^j(t,x)v \;=\; b_j^p(t,x)D_pv, \qquad \hat{\mathcal{M}}_n^j(t,x)v \;=\; b_{nj}^p(t,x)D_pv,$$

$j = 1,2,\dots,r$ (we sum with respect to p from 0 to d).

For all $p = 0,1,\dots,d$, $j,l = 1,2,\dots,r$ set

$$\begin{aligned}
\beta_j^{p(0)}(t,x,y) &= \Big(\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial y_h^2}\Big)\beta_j^p(t,x,y),\\
\beta_{nj}^{p(0)}(t,x,y) &= \frac{\partial}{\partial t}\beta_j^p(t,x,y),\\
\beta_j^{p(l)}(t,x,y) &= \frac{\partial}{\partial y_l}\beta_j^p(t,x,y),\\
\beta_{nj}^{p(l)}(t,x,y) &= \frac{\partial}{\partial y_l}\beta_j^p(t,x,y).
\end{aligned}$$

Now define

$$b_j^{p(l)}(t,x) \;=\; \beta_j^{p(l)}(t,x,y(t)), \qquad b_{nj}^{p(l)}(t,x) \;=\; \beta_{nj}^{p(l)}(t,x,y_n(t)),$$

$j = 1,2,\dots,r$, $l = 0,1,\dots,r$, $p = 0,1,\dots,d$, and then

$$\hat{\mathcal{M}}^{j(l)}(t,x)v \;=\; b_j^{p(l)}(t,x)D_pv, \qquad \hat{\mathcal{M}}_n^{j(l)}(t,x)v \;=\; b_{nj}^{p(l)}(t,x)D_pv,$$

$j = 1,2,\dots,r$, $l = 0,1,\dots,r$ (we sum with respect to p from 0 to d).

At last, for all $p = 0,1,\dots,d$, $j,k,l = 1,2,\dots,r$ set

$$\beta_j^{p(k0)}(t,x,y) \;=\; \Big(\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial y_h^2}\Big)\frac{\partial}{\partial y_k}\beta_j^p(t,x,y), \quad \beta_j^{p(kl)}(t,x,y) \;=\; \frac{\partial^2}{\partial y_l \partial y_k}\beta_j^p(t,x,y),$$

and define

$$b_j^{p(kl)}(t,x) \;=\; \beta_j^{p(kl)}(t,x,y_n(t)),$$

$j,k = 1,2,\dots,r$, $l = 0,1,\dots,r$, $p = 0,1,\dots,d$, and then

$$\hat{\mathcal{M}}^{j(kl)}(t,x)v \;=\; b_j^{p(kl)}(t,x)D_pv,$$

$j,k = 1,2,\dots,r$, $l = 0,1,\dots,r$ (we sum with respect to p from 0 to d).

Assigning

$$\begin{aligned}
\mathcal{L} &= \hat{\mathcal{L}} - \frac{1}{2}\hat{\mathcal{M}}^j\hat{\mathcal{M}}^j - \frac{1}{2}\hat{\mathcal{M}}^{j(j)}, \quad \mathcal{M}^j = \hat{\mathcal{M}}^j, \quad \mathcal{M}^{j(l)} = \hat{\mathcal{M}}^{j(l)},\\
&\qquad\qquad \mathcal{M}^{j(kl)} = \hat{\mathcal{M}}^{j(kl)},\\
\mathcal{L}_n &= \hat{\mathcal{L}}_n - \frac{1}{2}\hat{\mathcal{M}}_n^j\hat{\mathcal{M}}_n^j - \frac{1}{2}\hat{\mathcal{M}}_n^{j(j)}, \quad \mathcal{M}_n^j = \hat{\mathcal{M}}_n^j, \quad \mathcal{M}_n^{j(l)} = \hat{\mathcal{M}}_n^{j(l)}
\end{aligned}$$

we reduce problem (5.2.3)-(5.2.4) to the form

$$\begin{aligned} d\varphi(t,x) &= \mathcal{L}\varphi(t,x)dt + \mathcal{M}^j\varphi(t,x)dy^j(t) \\ &\quad + \frac{1}{2}\mathcal{M}^j\mathcal{M}^j\varphi(t,x)dt + \frac{1}{2}\mathcal{M}^{j(j)}\varphi(t,x)dt, \qquad (5.3.1) \\ \varphi(0,x) &= p_0(x), \qquad (5.3.2) \end{aligned}$$

equivalent to the "Stratonovich" equation (4.3.5), and we reduce problem (5.2.5)-(5.2.6) to the form

$$\begin{aligned} d\varphi_n(t,x) &= \mathcal{L}_n\varphi_n(t,x)dt + \mathcal{M}_n^j\varphi_n(t,x)dy_n^j(t), \qquad (5.3.3) \\ \varphi_n(0,x) &= p_{n0}(x), \qquad (5.3.4) \end{aligned}$$

which is equivalent to the partial differential equation (4.3.6). Our aim now is to show that in this case Assumptions 4.3.1-4.3.5 are satisfied. Indeed, by the Itô formula the processes $b_j^p(t,x)$, $b_{nj}^p(t,x)$, $b_j^{p(k)}(t,x)$ for all $j,k = 1,2,\dots,r$, $p = 0,1,\dots,d$ have stochastic differentials

$$\begin{aligned} db_j^p(t,x) &= b_j^{p(0)}(t,x)dt + b_j^{p(l)}(t,x)dW^l(t), \\ db_{nj}^p(t,x) &= b_{nj}^{p(0)}(t,x)dt + b_{nj}^{p(l)}(t,x)dW_n^l(t), \\ db_j^{p(k)}(t,x) &= b_j^{p(k0)}(t,x)dt + b_j^{p(kl)}(t,x)dW^l(t), \end{aligned}$$

respectively, on $[0,T]$ for all $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$. By a Fubini-type theorem for stochastic integrals (see p. 116, [13]) for every $\psi \in C_0^\infty(\mathbb{R}^d)$ the stochastic processes $(b_j^p(t),\psi)_0$, $(b_{nj}^p(t),\psi)_0$, $(b_j^{p(k)}(t),\psi)_0$ have stochastic differentials

$$\begin{aligned} d(b_j^p(t),\psi)_0 &= (b_j^{p(0)}(t),\psi)_0dt + (b_j^{p(l)}(t),\psi)_0dW^l(t), \\ d(b_{nj}^p(t),\psi)_0 &= (b_{nj}^{p(0)}(t),\psi)_0dt + (b_{nj}^{p(l)}(t),\psi)_0dW_n^l(t), \\ d(b_j^{p(k)}(t),\psi)_0 &= (b_j^{p(k0)}(t),\psi)_0dt + (b_j^{p(kl)}(t),\psi)_0dW^l(t), \end{aligned}$$

respectively, on $[0,T]$ for all $n \in \mathbb{N}$. It is easy to show that Assumptions 4.3.3, 4.3.4 hold.

Finally, if H^j is uniformly bounded in (t,x,y) for every $j = 1,2,\dots,r$, then by Girsanov's Theorem, p. 207, [17], we can introduce on the measurable space $(\Omega,\mathcal{F})$ a new probability measure $\tilde{\mathbf{P}}$ by

$$\tilde{\mathbf{P}}(d\omega) = \rho_T(H)(\omega)\mathbf{P}(d\omega),$$

such that $(\Omega,\mathcal{F},\tilde{\mathbf{P}})$ is a probability space and $W(t) + \int_0^t H(s,x(s),y(s))ds$ is an r-dimensional Wiener process on the interval $[0,T]$ on $(\Omega,\mathcal{F},\tilde{\mathbf{P}})$. Above $\rho_T(H)$, called the exponential martingale, is defined by

$$\rho_T(H) = \exp\{-\sum_j \int_0^T H^j(s,x(s),y(s))dW^j(s)$$
$$-\frac{1}{2}\sum_j \int_0^T (H^j(s,x(s),y(s)))^2 ds\}.$$

Notice, that condition $E\rho_T(H)=1$ of the Girsanov's theorem holds by Lemma 1, p. 204, [17]. Therefore, Assumption 5.2.1 implies Assumption 4.3.1. We apply Theorem 4.3.1.

□

5.3.2 Convergence $p_n \to p$

We have, for every $t \in [0,T]$

$$|p(t)-p_n(t)|_{W_2^m} \le \frac{|\varphi(t)|_{W_2^m}|(\varphi(t),1)_{L_2}-(\varphi_n(t),1)_{L_2}|}{|(\varphi(t),1)_{L_2}(\varphi_n(t),1)_{L_2}|} + \frac{|\varphi(t)-\varphi_n(t)|_{W_2^m}}{(\varphi_n(t),1)_{L_2}}.$$

Therefore, since by Theorem 5.2.1 almost surely

$$\inf_{t\le T}(\varphi(t),1)_{L_2} > \zeta \tag{5.3.5}$$

for some positive random variable ζ, the problem of estimation of $|p-p_n|_{W_2^m}$, $\int_0^T |p-p_n|^2_{W_2^{m+1}}ds$ can be reduced to the estimation of $|(\varphi,1)_{L_2}-(\varphi_n,1)_{L_2}|$. Namely, in order to prove the second part of the theorem it is enough to show that

$$(\varphi,1)_{L_2}-(\varphi_n,1)_{L_2} = O(n^{-\kappa}) \tag{5.3.6}$$

for all $\kappa < \alpha/2$. Then (5.3.5), (5.3.6) imply also

$$\inf_{t\le T}(\varphi_n(t),1)_{L_2} > \zeta_1$$

almost surely for some positive random variable ζ_1 and for sufficiently large n. It is clear that this is what we need to get our results from the first inequality in the section.

Therefore, it suffices to show that (5.3.6) holds for any $\kappa < \alpha$. Introduce function $\rho(x) = (1+x^2)^{\gamma/2}$. Then

$$|(\varphi(t),1)_{L_2}-(\varphi_n(t),1)_{L_2}| \le |(\rho(\varphi(t)-\varphi_n(t)),\rho^{-1})_{L_2}| \le C|\rho(\varphi(t)-\varphi_n(t))|_{L_2}$$

for every $t \in [0,T]$, where $C = |\rho^{-1}|_{L_2}$ is finite for $\gamma > d$, and therefore it is enough to verify that

$$|\rho(\varphi-\varphi_n)|_{L_2} = O(n^{-\kappa}) \tag{5.3.7}$$

for all $\kappa < \alpha/2$.

It follows from problem (5.3.1)-(5.3.2) that for function $\rho(x)$ defined above $\varphi(t,x)$ satisfies equation

$$\begin{aligned} d(\rho(x)\varphi(t,x)) &= \rho(x)\mathcal{L}\varphi(t,x)dt + \rho(x)\mathcal{M}^j\varphi(t,x)dy^j(t) \\ &+ \frac{1}{2}\rho(x)\mathcal{M}^j\mathcal{M}^j\varphi(t,x)dt + \frac{1}{2}\rho(x)\mathcal{M}^{j(j)}\varphi(t,x)dt \end{aligned} \tag{5.3.8}$$

with the initial condition

$$\varphi(0,x) = p_0(x).$$

Introducing

$$\psi(t,x) = \rho(x)\varphi(t,x)$$

we can rewrite this in the form

$$\begin{aligned} d(\psi(t,x)) &= L(\psi(t,x))dt + M^j(\psi(t,x))dy^j \\ &+ \frac{1}{2}M^jM^j(\psi(t,x))dt + \frac{1}{2}M^{j(j)}(\psi(t,x))dt \end{aligned} \tag{5.3.9}$$

with the initial condition

$$\psi(0,x) = \rho(x)p_0(x), \tag{5.3.10}$$

where

$$Lv = D_p(\mathbf{a}^{pq}(t,x)D_qv), \quad M^jv = \mathbf{b}_j^p(t,x)D_pv \quad M^{j(l)}v = \mathbf{b}_j^{p(l)}(t,x)D_pv,$$

and coefficients $\mathbf{a}^{pq}$, $\mathbf{b}_j^p$, $\mathbf{b}_j^{p(l)}$ will be written below in terms of a^{pq}, b_j^p, $b_j^{p(l)}$. Similarly, introducing

$$\psi_n(t,x) = \rho(x)\varphi_n(t,x),$$

we get that $\psi_n(t,x)$ satisfies for all $n \in \mathbb{N}$ equation

$$d(\psi_n(t,x)) = L_n(\psi_n(t,x))dt + M_n^j(\psi_n(t,x))dy_n^j \tag{5.3.11}$$

with the initial condition

$$\psi_n(0,x) = \rho(x)p_{n0}(x). \tag{5.3.12}$$

Operators L_n, M_n^j, $M_n^{j(l)}$ are defined by

$$L_nv = D_p(\mathbf{a}_n^{pq}(t,x)D_qv), \quad M_n^jv = \mathbf{b}_{nj}^p(t,x)D_pv, \quad M_n^{j(l)}v = \mathbf{b}_{nj}^{p(l)}(t,x)D_pv,$$

coefficients $\mathbf{a}_n^{pq}$, $\mathbf{b}_{nj}^p$, $\mathbf{b}_{nj}^{p(l)}$ will be written below in terms of a_n^{pq}, b_{nj}^p, $b_{nj}^{p(l)}$. Below we verify that operators L, L_n, M^j, M_n^j satisfy Assumptions 4.3.2-4.3.5 and apply Theorem 4.3.1.

For simplicity of notations we will omit parameters t, x. It is easy to show that

$$\rho \mathcal{M}^j v \;=\; \sum_{p=0}^{d} b_j^p D_p(\rho v) \;-\; \rho v \sum_{p=1}^{d} b_j^p \mu_p,$$

where $\mu_p = D_p \rho(x)/\rho(x)$. Denoting

$$\begin{aligned} \mathbf{b}_j^0 &= b_j^0 - \sum_{p=1}^{d} b_j^p \mu_p, \\ \mathbf{b}_j^p &= b_j^p, \qquad p \neq 0 \end{aligned}$$

we get

$$\rho \mathcal{M}^j v \;=\; M^j(\rho v), \qquad j = 1, 2, \ldots, r. \tag{5.3.13}$$

Set $u = \mathcal{M}^l v$ and apply the previous formula to $\rho \mathcal{M}^j u$ twice. We get

$$\rho \mathcal{M}^j \mathcal{M}^l v \;=\; M^j M^l(\rho v), \qquad j, l = 1, 2, \ldots, r. \tag{5.3.14}$$

Denoting

$$\begin{aligned} \mathbf{b}_j^{0(l)} &= b_j^{0(l)} - \sum_{p=1}^{d} b_j^{p(l)} \mu_p, \\ \mathbf{b}_j^{p(l)} &= b_j^{p(l)}, \qquad p \neq 0 \end{aligned}$$

we show that

$$\rho \mathcal{M}^{j(l)} v \;=\; M^{j(l)}(\rho v), \qquad j = 1, 2, \ldots, r, \; l = 0, 1, \ldots, r. \tag{5.3.15}$$

In order to get a similar relation for $\mathcal{L}$ and L we introduce intermediate operators $\mathcal{N}^p$, N^p defined by

$$\mathcal{N}^p v \;=\; \sum_{q=0}^{d} a^{pq} D_q v, \qquad N^p v \;=\; \sum_{q=0}^{d} \mathbf{c}^{pq} D_q v,$$

so that

$$\mathcal{L} v \;=\; \sum_{p=0}^{d} D_p \mathcal{N}^p v, \qquad L v \;=\; \sum_{p=0}^{d} D_p N^p v.$$

As above assigning

$$\begin{aligned} \mathbf{c}^{p0} &= a^{p0} - \sum_{q=1}^{d} a^{pq} \mu_q, \\ \mathbf{c}^{pq} &= a^{pq}, \qquad q \neq 0 \end{aligned}$$

we get relation

$$\rho \mathcal{N}^p v \;=\; N^p(\rho v), \qquad p = 0, 1, \ldots, r.$$

Then

$$\begin{aligned}
\rho\mathcal{L}v &= \sum_{p=0}^{d} \rho D_p \mathcal{N}^p v \\
&= \sum_{p=0}^{d} D_p(\rho\mathcal{N}^p v) - \sum_{p=1}^{d} (\rho\mathcal{N}^p v)\mu_p \\
&= \sum_{p=0}^{d} D_p(N^p(\rho v)) - \sum_{p=1}^{d} N^p(\rho v)\mu_p,
\end{aligned}$$

Therefore, defining

$$\begin{aligned}
\mathbf{a}^{0q} &= \mathbf{c}^{0q} - \sum_{p=1}^{d} \mathbf{c}^{pq}\mu_p, \\
\mathbf{a}^{pq} &= \mathbf{c}^{pq}, \qquad p \neq 0
\end{aligned}$$

we get

$$\rho\mathcal{L}v = L(\rho v). \tag{5.3.16}$$

Hence, using relations (5.3.13)-(5.3.16) we indeed reduce equation (5.3.8) to equation (5.3.9).

Similarly to $\mathbf{a}^{pq}$, $\mathbf{b}^p_j$, $\mathbf{b}^{p(l)}_j$ we can define coefficients $\mathbf{a}^{pq}_n$, $\mathbf{b}^p_{nj}$, $\mathbf{b}^{p(l)}_{nj}$. It is easy to show that if a^{pq}, b^p_j, ..., $b^{p(l)}_{nj}$ satisfy Assumptions 4.3.2-4.3.5 then so do $\mathbf{a}^{pq}$, $\mathbf{b}^p_j$, ..., $\mathbf{b}^{p(l)}_{nj}$. Applying Theorem 4.3.1 to problems (5.3.9)-(5.3.10), (5.3.11)-(5.3.12) we get

$$|\psi - \psi_n|_{L_2} = O(n^{-\kappa})$$

for every $\kappa < \alpha/2$, which leads to (5.3.7).

$\square$

Chapter 6

Examples of Wiener Process Approximations

6.1 Introduction

The chapters above were devoted to the study of the approximations of stochastic differential equations (SDEs) by ordinary differential equations (ODEs). Namely, we investigated the rate of the convergence of the solutions for the approximating ODEs to the solution for the original SDEs.

Initially we considered a "Stratonovich" SDE. We replaced the Wiener process W in this equation by an approximation system of stochastic processes $\{W_n\}_{n\in\mathbb{N}}$. In this chapter we study the two most common types of these approximations for the Wiener process W: polygonal approximation and smoothing. We verify that in these two cases sequences $\{W_n\}_{n\in\mathbb{N}}$ satisfy Assumption 2.3.1.

Assumption 6.1.1. *There exists a positive number α such that for every $\kappa < \alpha$ and every positive δ*

$$\begin{aligned} &(1) \quad W - W_n \;=\; O(n^{-\kappa}), \\ &(2) \quad S_n \;=\; O(n^{-\kappa}), \\ &(3) \quad ||S_n|| \;=\; O(\ln^{\delta} n). \end{aligned}$$

Here we used the notation

$$S_n^{jl}(t) \;=\; \int_0^t (W^j(s) - W_n^j(s))dW_n^l(s) \;-\; \frac{1}{2}\delta_{jl}t,$$

where δ_{jl} is the Kronecker's symbol which assumes 1 if $j = l$, and 0 otherwise.

6.2 Preliminaries

In this section we list some statements that we will use later in the chapter.

Theorem 6.2.1. *Let W be a Wiener process and let $\varepsilon \in (0, 1/2)$, $T \in (0, \infty)$. Then there exists a random variable $N = N(\omega)$, depending also on ε, T, such that $EN^p < \infty$ for all $p \in [0, \infty)$, and for any $\omega \in \Omega$*

$$|W_t - W_s| \;\leq\; N(\omega)|t - s|^{\frac{1}{2} - \varepsilon}$$

for all $t, s \in [0, T]$.

This is a well known fact. The proof can be found in [16], p. 36. The following lemma is a corollary of Lemma 2.4.2.

Lemma 6.2.2. *Let $\{\xi_n\}_{n \in \mathbb{N}}$ be a sequence of real random variables, and let β be a positive number. Suppose that for every $n \in \mathbb{N}$ and every positive integer r*

$$(E|\xi_n|^{2r})^{1/2r} \;\leq\; c_\beta n^{-\beta},$$

where c_β may depend on r but not on n. Then for every positive $\gamma < \beta$

$$\xi_n \;=\; O(n^{-\gamma}).$$

Proof. We fix any value of γ, then set any $r > \frac{1}{2}(\beta - \gamma)^{-1}$, and apply Lemma 2.4.2. □

Lemma 6.2.3. *Let $\xi(t)$, $t \in [0, T]$ be simultaneously a continuous martingale and a Gaussian process with mean 0 and variance σ_t. Then*

$$\mathbf{P}\{\sup_{t \leq T} \xi(t) \geq a\} \;\leq\; \exp\left(-\frac{1}{2}\frac{a^2}{\sigma_T}\right). \tag{6.2.1}$$

Proof. By properties of positive submartingales (see Theorem 3.8, [17]) and normal random variables for arbitrary λ

$$\mathbf{P}\{\sup_{t \leq T} \xi(t) \geq a\} \;\leq\; \mathbf{P}\{\sup_{t \leq T} e^{\lambda \xi(t)} \geq e^{\lambda a}\} \;\leq\; \frac{1}{e^{\lambda a}} E e^{\lambda \xi(T)} \;=\; \frac{1}{e^{\lambda a}} e^{\frac{1}{2}\lambda^2 \sigma_T}.$$

Minimizing the last term with respect to λ we derive (6.2.1). □

Lemma 6.2.4. *Let stochastic processes $M(t)$ simultaneously be a continuous martingale and a Gaussian process with mean 0 and variance $\sigma_t = \alpha t$, where α is a positive constant. Denote*

$$X \;=\; \sup_{0 < t \leq e^{-2}} \frac{M(t)}{\sqrt{t \ln \ln \frac{1}{t}}}.$$

Then $EX^p < \infty$ for any integer $p \geq 2$.

Proof. Note, that the interval $(0, e^{-2}]$ is dictated by function $\sqrt{t \ln\ln(1/t)}$. Introduce a partition $\{t_k\}_{k=2,3,\dots}$ of the interval $(0, e^{-2}]$, $t_k = e^{-k}$. Then

$$X = \sup_{k\geq 2} \sup_{t_{k+1}<t\leq t_k} \frac{M(t)}{\sqrt{t \ln\ln \frac{1}{t}}}.$$

Let us show that for every $z \geq 0$

$$\mathbf{P}\{X \geq z\} \leq \sum_{k=2}^{\infty} e^{-cz^2 \ln(1+k)} \tag{6.2.2}$$

for some positive constant c. Indeed, we have,

$$\begin{aligned} \mathbf{P}\{X \geq z\} &\leq \sum_{k=2}^{\infty} \mathbf{P}\left\{ \sup_{t_{k+1}<t\leq t_k} \frac{M(t)}{\sqrt{t\ln\ln\frac{1}{t}}} \geq z \right\} \\ &\leq \sum_{k=2}^{\infty} \mathbf{P}\left\{ \sup_{0<t\leq t_k} M(t) \geq z\sqrt{t_{k+1}\ln\ln\frac{1}{t_{k+1}}} \right\}, \end{aligned}$$

and by Lemma 6.2.3

$$\mathbf{P}\{X \geq z\} \leq \sum_{k=2}^{\infty} \exp\left(-\frac{1}{2}\frac{z^2 t_{k+1}\ln\ln\frac{1}{t_{k+1}}}{\sigma_{t_k}} \right) = \sum_{k=2}^{\infty} \exp\left(-\frac{1}{2}\frac{z^2\ln(1+k)}{\alpha e} \right),$$

which proves (6.2.2).

Let us separate large values of X, so that $X_N = X\chi_{X\geq N}$, and estimate EX_N^p. The random variable X_N^p is non-negative. Therefore,

$$EX_N^p = \int_0^\infty \mathbf{P}\{X_N^p \geq x\}dx = \int_N^\infty \mathbf{P}\{X \geq x^{1/p}\}dx.$$

Using estimate (6.2.2) we get,

$$\begin{aligned} EX_N^p &\leq \sum_{k=2}^{\infty} \int_N^\infty e^{-cx^{2/p}\ln(1+k)}dx = \sum_{k=2}^{\infty} \int_0^\infty e^{-c(x+N)^{2/p}\ln(1+k)}dx \\ &\leq \sum_{k=2}^{\infty} \int_0^\infty e^{-c(p)(x^{2/p}+N^{2/p})\ln(1+k)}dx \end{aligned}$$

for some constant $c(p)$ depending only on p. Then

$$EX_N^p \leq d(p)\sum_{k=2}^{\infty} e^{-c(p)N^{2/p}\ln(1+k)} = d(p)\sum_{k=2}^{\infty}(1+k)^{-c(p)N^{2/p}},$$

where $d(p) = \int_0^\infty e^{-c(p)x^{2/p}\ln 3}dx$ is finite for any $p \geq 2$. Taking N large enough such that the sum is finite we get $EX_N^p < \infty$.

Finally,

$$EX^p = EX^p\chi_{X<N} + EX^p\chi_{X\geq N} \leq N^p + EX_N^p < \infty.$$

$\square$

6.3 Polygonal approximation

For simplicity assume that T is an integer. Let us define for any $n = 1, 2, \dots$ partition $\{t_k^n\}_{k=0}^{nT}$ of interval $[0, T]$, where $t_k^n = \frac{k}{n}$. For short we omit index n and use notation t_k. The polygonal approximation of the Wiener process W may be defined as $\widetilde{W}_n(t) = W(t_k) + n(t - t_k)(W(t_{k+1}) - W(t_k))$ in the interval $[t_k, t_{k+1})$ for $k = 0, 1, \dots, nT - 1$. However, this family is $\mathcal{F}_{t+\frac{1}{n}}$-adapted, and, consequently, not $\mathcal{F}_t$-adapted. Introducing a shifted $\mathcal{F}_t$-adapted version $W_n(t) = \widetilde{W}_n(t - \frac{1}{n})$ we define a natural approximation, i.e. an approximation which does not depend on the "future".

Definition 6.3.1. *We say that a sequence $\{W_n\}_{n \in \mathbb{N}}$ is a polygonal approximation of a Wiener processes W if it is defined as*

$$W_n(t) = W(t_{k-1}) + n(t - t_k)(W(t_k) - W(t_{k-1}))$$

for values of t from the interval $[t_k, t_{k+1})$ for $k = 0, 1, \dots, nT - 1$. To make the definition consistent we assume that $W(t)$ vanishes for negative t.

The following is the main statement of the section.

Proposition 6.3.1. *In the case of polygonal approximation Assumption 6.1.1 holds with $\alpha = 1/2$.*

We will need the following lemma.

Lemma 6.3.2. *For any integer $n \geq 1$ let $X_{n1}, \dots, X_{nn}$ be independent random variables such that for $k = 1, 2, \dots, n$ $EX_{nk} = 0$, and for a sufficiently large integer $p \geq 1$*

$$E|X_{nk}|^{2p} \leq c_p n^{-2p\beta},$$

for all $n \geq 1$ and some $\beta \in (\frac{1}{2}, 1]$, where c_p is a constant depending only on p. Then $S_n = \sum_{k=1}^n X_{nk}$ converges to 0 almost surely. Moreover, for every $\gamma < \beta - \frac{1}{2}$ there exists a finite random variable ξ_γ such that almost surely

$$|S_n| \leq \xi_\gamma n^{-\gamma}$$

for all $n \geq 1$.

Proof. First, show that for any multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$ such that $\alpha_1 + \dots + \alpha_n = 2p$ the following inequality holds:

$$E|X_{n1}^{\alpha_1} \dots X_{nn}^{\alpha_n}| \leq c_p n^{-2p\beta}.$$

Indeed, by Hölder's inequality

$$\begin{aligned} E|X_{n1}^{\alpha_1}\dots X_{nn}^{\alpha_n}| &\le \left(E|X_{n1}|^{2p}\right)^{\frac{\alpha_1}{2p}}\dots\left(E|X_{nn}|^{2p}\right)^{\frac{\alpha_n}{2p}} \\ &\le \prod_{k=1}^{n}\left(c_p n^{-2p\beta}\right)^{\frac{\alpha_k}{2p}} = c_p n^{-2p\beta}. \end{aligned}$$

Hence,

$$E(S_n)^{2p} \le (2p)!N(n,p)n^{-2p\beta},$$

where $N(n,p)$ is the number of those multi-indices $\alpha = (\alpha_1,\dots,\alpha_n)$ satisfying the conditions $\alpha_j \ne 1$ for all $j = 1,2,\dots,n$, and $\alpha_1 + \dots + \alpha_n = 2p$. Our aim is to show that $N(n,p) \le c(p)n^p$ for some constant $c(p)$ depending only on p.

The number of $\alpha_1,\dots,\alpha_n$ different from 0 is less than or equal to p. So $N(n,p)$ can be estimated by $\binom{n}{p}N(p)$. We take $\binom{n}{p}$ choices to place p elements to n positions. And for every such a choice we determine number $N(p)$ of all possibilities to choose p multi-indices $(\alpha_{i_1},\dots,\alpha_{i_p})$ for which $\alpha_{i_1}+\dots+\alpha_{i_p} = 2p$. Obviously $N(p)$ exceeds the number of all multi-indices with the elements different from 1.

Next, $N(p)$ equals to the number of multi-indices $(\beta_1,\dots,\beta_p)$ possessing the following properties: $\beta_j \ge 1$ for all $j = 1,2,\dots,p$, and $\beta_1 + \dots + \beta_p = 3p$. And, hence, is equal to the number of possibilities of placing $p-1$ elements to $3p-1$ positions. So,

$$N(n,p) \le \binom{n}{p}\binom{3p-1}{p-1} \le c(p)n^p,$$

where constant $c(p) = \frac{1}{p!}\binom{3p-1}{p-1}$ depends only on p, and

$$E(S_n)^{2p} \le c(p)n^{p(1-2\beta)}.$$

Lemma 6.2.2 ends the proof of the lemma. □

Proof of Proposition 6.3.1. Using Theorem 6.2.1 for every $\varepsilon > 0$ and every $t \in [t_k, t_{k+1})$ almost surely

$$|W^j(t) - W^j_n(t)| \le 2 \sup_{\substack{t,s\le T \\ |t-s|\le\frac{2}{n}}} |W^j(t) - W^j(s)| \le \zeta_\varepsilon n^{-\frac{1}{2}+\varepsilon},$$

where ζ_ε is a finite random variable depending only on ε. Assumption 6.1.1(1) holds. Next we verify Assumption 6.1.1(2), i.e. $|S_n^{jl}| = O(n^{-\gamma})$ for every j, l, and any $\gamma < 1/2$. We start with the situation where j and l coincide. By the Itô's formula

$$|W^j(t) - W^j_n(t)|^2 = 2\int_0^t (W^j(s) - W^j_n(s))d(W^j(s) - W^j_n(s)) + t,$$

and, hence,

$$S_n^{jj}(t) = \int_0^t (W^j(s) - W_n^j(s))dW^j(s) - \frac{1}{2}|W^j(t) - W_n^j(t)|^2. \tag{6.3.1}$$

By Burkholder-Davis-Gundy and Jensen's inequalities for any $p \geq 1$

$$E \sup_{t \leq T} |S_n^{jj}|^p \leq cE \sup_{\substack{t \leq T \\ |t-s| \leq \frac{2}{n}}} |W^j(t) - W^j(s)|^p \leq cn^{-p/2},$$

where constant c does not depend on n. It suffices to apply Lemma 6.2.2. Next, we verify Assumption 6.1.1(2) in the case of distinct j and l. By Definition 6.3.1 of the polygonal approximation

$$\begin{aligned} S_n^{jl}(t) &= n \sum_{k=0}^{nT-1} (W^l(t_k) - W^l(t_{k-1})) \int_{t_k}^{t_{k+1}} (W^j(s) - W^j(t_k))\chi_{s \leq t} ds \\ &+ n \sum_{k=0}^{nT-1} (W^l(t_k) - W^l(t_{k-1}))(W^j(t_k) - W^j(t_{k-1})) \\ &\times \int_{t_k}^{t_{k+1}} (n(s - t_k) + 1)\chi_{s \leq t} ds. \end{aligned}$$

Assume that for $n \in \mathbb{N}$ time parameter t belongs to the interval $[t_{m_n}, t_{m_n+1})$. Then

$$\begin{aligned} S_n^{jl}(t) &\leq \sum_{k=0}^{m_n} (W^l(t_k) - W^l(t_{k-1})) \int_0^{a_k^n} (W^j(t_k + \frac{u}{n}) - W^j(t_k))du \\ &+ \frac{5}{2} \sum_{k=0}^{m_n} (W^l(t_k) - W^l(t_{k-1}))(W^j(t_k) - W^j(t_{k-1})), \end{aligned}$$

where $a_k^n = n(t - t_{m_n})$ if $k = m_n$, and $a_k^n = 1$ otherwise. For $k \leq m_n$ denote the terms in the first sum of the last inequality as X_{nk}^{jl}, and the terms in the second sum as Y_{nk}^{jl}. Define $X_{nk}^{jl} = Y_{nk}^{jl} = 0$ for $k > m_n$. Then

$$S_n^{jl}(t) \leq \sum_{k=0}^{nT-1} X_{nk}^{jl} + \frac{5}{2} \sum_{k=0}^{nT-1} Y_{nk}^{jl}.$$

It is easy to show that for every k, $EX_{nk}^{jl} = EY_{nk}^{jl} = 0$, and

$$\begin{aligned} E|X_{nk}^{jl}|^{2p} &\leq E|Z_{\frac{1}{n}}|^{2p} \int_0^{a_k^n} E|Z_{\frac{u}{n}}|^{2p} du \leq \frac{((2p-1)!!)^2}{p+1} n^{-2p}, \\ E|Y_{nk}^{jl}|^{2p} &\leq \left(E|Z_{\frac{1}{n}}|^{2p}\right)^2 = ((2p-1)!!)^2 n^{-2p}, \end{aligned}$$

where Z_v is a Gaussian random variable with mean 0 and variance v. By Lemma 6.3.2, $|S_n^{jl}| = O(n^{-\gamma})$ for any $\gamma < \frac{1}{2}$. Hence, Assumption 6.1.1(2) is satisfied with $\alpha = 1/2$.

Now we go on to Assumption 6.1.1(3). We have

Now we go on to Assumption 6.1.1(3). We have Now we go on to Assumption 6.1.1(3). We have Now we go on to Assumption 6.1.1(3). We have Now we go on to Assumption 6.1.1(3). We have Now we go on to Assumption 6.1.1(3). We have

$$
\begin{aligned}
&||S_n||(T) \\
&\leq n \sum_{j,l} \int_{t_k}^{t_{k+1}} \sum_{k=0}^{nT-1} |(W^j(t) - W^j(t_{k-1})) - n(t-t_k)(W^j(t_k) - W^j(t_{k-1}))| \\
&\quad \times |W^l(t_k) - W^l(t_{k-1})| dt \ + \ \frac{m}{2} T \\
&\leq \sum_{j,l} \int_0^1 \left\{ \sum_{k=0}^{nT-1} |W^j(t_k + \frac{u}{n}) - W^j(t_k)||W^l(t_k) - W^l(t_{k-1})| du \right\} \\
&\quad + \sum_{j,l} \sum_{k=0}^{nT-1} |W^j(t_k) - W^j(t_{k-1})||W^l(t_k) - W^l(t_{k-1})| \ + \ \frac{m}{2} T \\
&= \sum_{j,l} \int_0^1 \left\{ \sum_{k=0}^{nT-1} X_{nk}^{jl} \right\} du + \sum_{j,l} \sum_{k=0}^{nT-1} Y_{nk}^{jl} + C,
\end{aligned}
$$

$$
\begin{aligned}
&||S_n||(T) \\
&\leq \ n \sum_{j,l} \int_{t_k}^{t_{k+1}} \sum_{k=0}^{nT-1} |(W^j(t) - W^j(t_{k-1})) \\
&\qquad\qquad\qquad -n(t-t_k)(W^j(t_k) - W^j(t_{k-1}))| \\
&\quad \times |W^l(t_k) - W^l(t_{k-1})| dt \ + \ \frac{m}{2} T \\
&\leq \ \sum_{j,l} \int_0^1 \left\{ \sum_{k=0}^{nT-1} |W^j(t_k + \frac{u}{n}) - W^j(t_k)||W^l(t_k) - W^l(t_{k-1})| du \right\} \\
&\quad + \sum_{j,l} \sum_{k=0}^{nT-1} |W^j(t_k) - W^j(t_{k-1})||W^l(t_k) - W^l(t_{k-1})| \ + \ \frac{m}{2} T \\
&= \ \sum_{j,l} \int_0^1 \left\{ \sum_{k=0}^{nT-1} X_{nk}^{jl} \right\} du + \sum_{j,l} \sum_{k=0}^{nT-1} Y_{nk}^{jl} + C,
\end{aligned}
$$

where

$$
\begin{aligned}
X_{nk}^{jl} &= |W^j(t_k + \frac{u}{n}) - W^j(t_k)||W^l(t_k) - W^l(t_{k-1})| \ - \ \mu_{\frac{u}{n}} \mu_{\frac{1}{n}}, \\
Y_{nk}^{jj} &= |W^j(t_k) - W^j(t_{k-1})|^2 \ - \ \frac{1}{n}, \\
Y_{nk}^{jl} &= |W^j(t_k) - W^j(t_{k-1})||W^l(t_k) - W^l(t_{k-1})| \ - \ \mu_{\frac{1}{n}}^2, \quad j \neq l,
\end{aligned}
$$

$$C = \frac{m}{2}T + \sum_{j,l}\sum_{i=0}^{nT-1}\int_0^1 \mu_{\frac{u}{n}}\mu_{\frac{1}{n}}du + \sum_{j}\sum_{k=0}^{nT-1}\frac{1}{n} + \sum_{j\neq l}\sum_{i=0}^{nT-1}\mu^2_{\frac{1}{n}}.$$

Here μ_v is the expectation of the module of a Gaussian random variable Z_v distributed with mean 0 and variation v. Obviously, $\mu_v^2 \leq v$ for any v, and hence

$$C \leq \frac{m}{2}T + 2m^2T.$$

Furthermore, we check conditions of Lemma 6.3.2 for X^{jl}_{nk} and Y^{jl}_{nk}. First of all, $EX^{jl}_{nk} = EY^{jl}_{nk} = 0$. Next,

$$E|Y^{jj}_{nk}|^{2p} = \sum_{i=0}^{2p}(-1)^{2p-i}\binom{2p}{i}\frac{1}{n^{2p-i}}E|Z_{\frac{1}{n}}|^{2i} \leq c_p n^{-2p},$$

where constant c_p depends only on p. The same sort of arguments shows that also Y^{jl}_{nk} for $j \neq l$ and X^{jl}_{nk} can be estimated by $c_p n^{-2p}$. By Lemma 6.3.2 both $\sum_k X^{jl}_{nk}$ and $\sum_k Y^{jl}_{nk}$ converge to 0 as n tends to infinity; $||S_n||(T)$, and hence $||S_n||(t)$ for any $t \leq T$, can be estimated by a finite random variable, i.e. $||S_n|| = O(1)$. This ends the proof of Proposition 6.3.1. □

6.4 Smoothing the Wiener process

In this section we study another type of approximation of the driving process.

Definition 6.4.1. We say that an approximation family $\{W_n\}_{n\in\mathbb{N}}$ is the *smoothing* of a Wiener process W if

$$W_n(t) = \int_0^1 W(t - \frac{u}{n})du.$$

Here we assume that $W(t)$ vanishes for negative t.

Proposition 6.4.1. *Assumption 6.1.1 holds when $W_n(t)$ is a smoothing of a Wiener process.*

We will need the following lemma.

Lemma 6.4.2. *Let W be a 1-dimensional Wiener process, $\{W_n\}_{n\in\mathbb{N}}$ its smoothing, $\{f_n\}_{n\in\mathbb{N}}$ a family of one dimensional stochastic processes independent of W. Assume that for some $\kappa > 0$ and positive integer p*

$$\left(E\sup_{t\leq T}|f_n(t)|^p\right)^{1/p} \leq c_0 n^{-\kappa}. \qquad (6.4.1)$$

Then

$$\left(E\sup_{t\leq T}\left|\int_0^t f_n(s)dW_n(s)\right|^p\right)^{1/p} \leq cn^{-\kappa} \qquad (6.4.2)$$

for some constant c independent of n, providing that the integral exists.

Proof. Changing the variable of integration in the definition of the smoothing, it is easy to show that

$$\frac{dW_n(t)}{dt} = n(W(t) - W(t-\frac{1}{n})) = n\int_{t-\frac{1}{n}}^{t} dW(v) \tag{6.4.3}$$

Define $W(t) = W(T)$, $f_n(t) = f_n(T)$ for $t > T$. Substituting (6.4.3) and changing the order of integration in $\int_0^t f_n(s)dW_n(s)$ we obtain the following estimate:

$$E\sup_{t\le T}\left|\int_0^t f_n(s)dW_n(s)\right|^p \le c(J_1 + J_2),$$

where

$$\begin{aligned}
J_1 &= c_p n^p E\sup_{t\le T}\left|\int_0^t \int_v^{v+\frac{1}{n}} f_n(s)dsdW(v)\right|^p, \\
J_2 &= c_p n^p E\sup_{t\le T}\left|\int_{t-\frac{1}{n}}^t \int_t^{v+\frac{1}{n}} f_n(s)dsdW(v)\right|^p,
\end{aligned}$$

constant c_p depends only on p. Applying Burkholder-Davis-Gundy, and then a sequence of Jensen's inequalities, and using conditions of the lemma, we get

$$J_1 \le cn^{-\kappa p}, \qquad J_2 \le cn^{-(\kappa+\frac{1}{2})p},$$

where constant c does not depend on n. □

Proof of Proposition 6.4.1. Assumption 6.1.1(1) holds by Theorem 6.2.1 since for every $\varepsilon > 0$ almost surely

$$\sup_{t\le T}|W^j(t) - W_n^j(t)| \le \sup_{\substack{t,u\le T\\ |t-u|<\frac{1}{n}}} |W^j(t) - W^j(u)| \le \zeta_\varepsilon n^{-\frac{1}{2}+\varepsilon}, \tag{6.4.4}$$

where ζ_ε is a finite random variable which does not depend on n.

Next, we verify Assumption 6.1.1(2) in the situation where j and l coincide. Using (6.3.1) we have

$$\begin{aligned}
E\sup_{t\le T}|S_n^{jj}|^p \le\; & cE\sup_{t\le T}\left|\int_0^t\int_0^1 (W^j(s) - W^j(s-\frac{u}{n}))du\, dW^j(s)\right|^p \\
& - cE\sup_{t\le T}\left|\int_0^1 (W^j(s) - W^j(s-\frac{u}{n}))du\right|^{2p}.
\end{aligned}$$

Then using Burkholder-Davis-Gundy and Jensen's inequalities

$$E\sup_{t\le T}|S_n^{jj}|^p \le cE\sup_{\substack{t\le T\\ |t-s|\le\frac{1}{n}}} |W^j(t) - W^j(s)|^p \le cn^{-p/2}$$

for any $p \geq 1$, where constant c does not depend on n. It suffices to apply Lemma 6.2.2. In the case $j \neq l$ we assign $f_n(t) = W^j(t) - W_n^j(t)$, $W(t) = W^l(t)$. Inequality (6.4.1) holds by (6.4.4). Assumption 6.1.1(2) is satisfied by Lemma 6.4.2 and Lemma 6.2.2.

It is remained to verify Assumption 6.1.1(3). By Definition 6.4.1

$$||S_n||(t) \leq n \sum_{j,l} \int_0^t \int_0^1 |W^j(s) - W^j(s - \frac{u}{n})| du |W^l(s) - W^l(s - \frac{1}{n})| ds + \frac{mt}{2}.$$

Define two random variables

$$\xi_{s,u}^j = \sup_{n \geq 3} \frac{|W^j(s) - W^j(s - u/n)|}{\sqrt{\frac{1}{n} \ln \ln n}} \quad \text{and} \quad \eta_s^l = \sup_{n \geq 3} \frac{|W^l(s) - W^l(s - 1/n)|}{\sqrt{\frac{1}{n} \ln \ln n}}$$

depending on parameters s, u. For fixed s, u they are equivalent to the random variables

$$\widetilde{\xi_{s,u}^j} = \sup_{n \geq 3} \frac{\sqrt{u} |\widetilde{W^j}(1/n)|}{\sqrt{\frac{1}{n} \ln \ln n}} \quad \text{and} \quad \widetilde{\eta_s^l} = \sup_{n \geq 3} \frac{|\bar{W}^l(1/n)|}{\sqrt{\frac{1}{n} \ln \ln n}},$$

respectively, where $\widetilde{W^j}(t)$ and $\bar{W}^l(t)$ are two Wiener processes depending on s, u, j, l. By Lemma 6.2.4

$$E \int_0^t \int_0^1 (\xi_{s,u}^j)^2 du ds = \int_0^t \int_0^1 E(\widetilde{\xi_{s,u}^j})^2 du ds < \infty$$

and

$$E \int_0^t (\eta_s^l)^2 ds = \int_0^t E(\widetilde{\eta_s^l})^2 ds < \infty,$$

and, consequently, almost surely

$$\int_0^t \int_0^1 (\xi_{s,u}^j)^2 du ds < \infty \quad \text{and} \quad \int_0^t (\eta_s^l)^2 ds < \infty.$$

Finally,

$$||S_n||(t) \leq \sum_{j,l} \int_0^t \int_0^1 \xi_{s,u}^j du \, \eta_s^l ds \ln \ln n \leq \zeta \ln \ln n,$$

where

$$\zeta = \sum_{j,l} \left(\int_0^t \int_0^1 (\xi_{s,u}^j)^2 du ds \right)^{1/2} \left(\int_0^t (\eta_s^l)^2 ds \right)^{1/2}$$

is an almost surely finite random variable. This implies for every positive δ

$$||S_n||(t) \leq \zeta_\delta \ln^\delta n$$

for some a.s. finite random variable ζ_δ. The proof of Proposition 6.4.1 is complete.

□

Bibliography

[1] Doss, H. *Liens entre équations différentielles stochastiques et ordinaires* (in French). Ann. Inst. H. Poincaré Sect. B (N.S.), 13 (1977), pp. 99-125.

[2] Faddeev, L. D.; Gorin, E. A.; Kostyuchenko, A. G.; Krasnosel'skiĭ, M. A.; Kreĭn, S. G.; Maslov, V. P.; Mityagin, B. S.; Petunin, Yu. I.; Rutitskii, Ya. B.; Sobolev, V. I.; Stetsenko, V. Ya.; Tsitlanadze, E. S.; Vilenkin, N. Ya. *Functional analysis.* 1972. Wolters-Noordhoff Publishing, Groningen.

[3] Fujisaki, M.; Kallianpur, G.; Kunita, H. *Stochastic differential equations for the nonlinear filtering problem.* Osaka J. Math., 9 (1972), pp. 19-40.

[4] Gikhman, I. I. and Skorokhod, A. V. *Introduction to the Theory of Random Processes.* 1969. Philadelphia: W.B. Saunders Co.

[5] Gyöngy, I. *On the approximation of stochastic differential equations.* Stochastics, 23 (1988), pp. 331-352.

[6] Gyöngy, I. *On the approximation of stochastic partial differential equations I.* Stochastics, 25 (1988), pp. 59-85.

[7] Gyöngy, I. *On the approximation of stochastic partial differential equations II.* Stochastics, 26 (1989), pp. 129-164.

[8] Gyöngy, I. *A note on Euler's approximations.* Potential Analysis, 8 (1998), pp. 205-216.

[9] Gyöngy, I.; Krylov, N. V. *On stochastic equations with respect to semimartingales II, Ito formula in Banach spaces.* Stochastics, 6 (1982), pp. 153-174.

[10] Gyöngy, I.; Shmatkov, A. *On the rate of convergence of Wong-Zakai approximations.* Preprint (2005).

[11] Gyöngy, I.; Shmatkov, A. *Rate of convergence of Wong-Zakai approximations for stochastic partial Differential Equations.* Preprint (2005).

[12] Ikeda, N.; Nakao, Sh.; Yamato, Yu. *A class of approximations of Brownian motion.* Publ. Res. Inst. Math. Sci., 13 (1977/78), no. 1, pp. 285-300.

[13] Ikeda, N.; Watanabe, S. *Stochastic differential equations and diffusion processes.* 1981. Amsterdam, Oxford, New York: North-Holland Publishing Co.

[14] Itô, K. *Stochastic differentials.* Appl. Math. Optim., 1 (1974/75), no. 4, pp. 374-381.

[15] Karatzas, I. and Shreve, S. E. *Brownian motion and stochastic calculus.* 1988. New York, London: Springer.

[16] Krylov, N. V. *Introduction to the theory of diffusion processes.* 1995. American Mathematical Society.

[17] Krylov, N. V. *Introduction to the theory of random processes.* 2002. American Mathematical Society.

[18] Krylov, N. V.; Rozovskiĭ, B.L. *Stochastic evolution equations.* (in Russian). Current problems in mathematics, 14 (1979), pp. 71-147.

[19] Krylov, N. V.; Rozovskiĭ, B.L. *Conditional distributions of diffusion processes.* (in Russian). Izv. Akad. Nauk SSSR Ser. Mat., 42 (1978), no. 2, pp. 356-378.

[20] Mackevicius, V. *S^p stability of symmetric stochastic differential equations.* Lietuvos Matematikos Rinkinys, 25 (1985), pp. 72-84. (in Russian)

[21] McShane, E.J. *Stochastic differential equations and models of random processes.* Proceedings of Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability theory, pp. 263-294. Univ. California Press, Berkeley, California, 1972.

[22] Rozovskiĭ, B. *Stochastic evolution systems. Linear theory and applications to nonlinear filtering.* 1990. Kluwer, Dordrecht.

[23] Wong, E. and Zakai, M. *On the convergence of ordinary integrals to stochastic integrals.* Ann. Math. Statist., 36 (1965), pp. 1560-1564.

[24] Wong, E. and Zakai, M. *On the relation between ordinary and stochastic differential equations.* Internat. J. Engrg. Sci., 3 (1965), pp. 213-229.

[25] Wong, E. and Zakai, M. *Riemann-Stieltjes approximations of stochastic integrals.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 12 (1969), pp. 87-97.